2013

CHARACTERIZATION OF G-PATCH MOTIF CONTRIBUTION TO PRP43 FUNCTION IN THE PRE-MESSENGER RNA SPLICING AND RIBOSOMAL RNA BIOGENESIS PATHWAYS

Daipayan Banerjee
University of Kentucky, daipayan.banerjee@gmail.com

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained and attached hereto needed written permission statements(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine).

I hereby grant to The University of Kentucky and its agents the non-exclusive license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless a preapproved embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's dissertation including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Daipayan Banerjee, Student

Dr. Brian C. Rymond, Major Professor

Dr. Brian C. Rymond, Director of Graduate Studies

CHARACTERIZATION OF G-PATCH MOTIF CONTRIBUTION TO PRP43 FUNCTION IN THE PRE-MESSENGER RNA SPLICING AND RIBOSOMAL RNA BIOGENESIS PATHWAYS

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Arts and Sciences at the University of Kentucky
by

Daipayan Banerjee

Lexington, Kentucky

Director: Dr. Brian C. Rymond, Linda and Jack Gill Professor of Biology,
Director of Graduate Studies

University of Kentucky
Lexington, KY 40506

2013

ABSTRACT OF DISSERTATION

CHARACTERIZATION OF G-PATCH MOTIF CONTRIBUTION TO PRP43 FUNCTION IN THE PRE-MESSENGER RNA SPLICING AND RIBOSOMAL RNA BIOGENESIS PATHWAYS

The DExD/H-box protein Prp43 is essential for two biological processes: nucleoplasmic pre-mRNA splicing and nucleolar rRNA maturation. The biological basis for the temporal and spatial regulation of Prp43 remains elusive. The Spp382/Ntr1, Sqs1/Pfa1 and Pxr1/Gno1 G-patch proteins bind to and activate the Prp43 DExD/H box-helicase in pre-mRNA splicing (Spp382) and rRNA processing (Sqs1, Pxr1). These Prp43-interacting proteins each contain the G-patch domain, a conserved sequence of ~48 amino acids that includes 6 highly conserved glycine (G) residues. Five annotated G-patch proteins in baker's yeast (i.e., Spp382, Pxr1, Spp2, Sqs1 and Ylr271) and with the possible exception of the uncharacterized Ylr271 protein, all are associated with ribonucleoprotein (RNP) complexes.

Understanding the role of G-patch proteins in modulating the DExD/H box protein Prp43 biological function was the motivation of this thesis. The G-patch domain has been proposed as a protein-protein or a protein-RNA interaction module for RNP proteins. This study found that the three Prp43-associated G-patch domains interact with Prp43 in a yeast 2 hybrid (Y2H) assay but differ in apparent relative affinities. Using a systemic Y2H analysis, I identified the conserved Winged-helix (WH) domain in Prp43 as a major binding site for G-patch motif. Intriguingly, removal of the non-essential N-terminal domain (NTD) of Prp43 (amino acids 2-94), greatly improves G-patch binding, suggesting that the NTD may play a role in modulating enzyme activity by the G-patch effectors. I identify a second site within the Pxr1 that strongly binds Prp43 but, unlike the G-patch, is dispensable for Pxr1 function *in vivo*.

By constructing chimeric proteins, I demonstrated that individual G-patch peptides differ in the ability to reconstitute Spp382 and Pxr1 function in support of pre-mRNA splicing and rRNA biogenesis, respectively. Through amino acid sequence comparisons and selective mutagenesis I identified several residues within the G-patch motif critical for Prp43-stimulated pre-mRNA splicing without greatly altering its ability to bind Prp43.

These data lead me to propose that the G-patch motif is not a simple Prp43 binding interface but may contribute more directly to substrate selection or Prp43 enzyme activation in the biologically distinct pre-mRNA splicing and rRNA processing pathways.

Key words: Prp43, DExD/H box helicase, G-patch domain, RNA splicing, ribosome biogenesis.

Daipayan Banerjee

Student's signature

April 12, 2013

CHARACTERIZATION OF G-PATCH MOTIF CONTRIBUTION TO PRP43 FUNCTION IN THE PRE-MESSENGER RNA SPLICING AND RIBOSOMAL RNA BIOGENESIS PATHWAYS

By

Daipayan Banerjee

Dr. Brian C. Rymond

Director of Dissertation

Dr. Brian C. Rymond

Director of Graduate Studies

April 12, 2013

This dissertation is dedicated to my father Mr. Rabindra Nath Banerjee, my mother Mrs. Smita Banerjee, my aunt Pushi (Puspa Ghosh) and my wife Swagata Ghosh.

ACKNOWLEDGEMENT

For me this dissertation is a life changing journey and became possible only with the support and love of many.

First and foremost, I would like to thank my advisor Dr. Brian Rymond for everything he has taught me, his highly valued guidance throughout the entire process and his continuous patience all along. He is a major catalyst in my transformation, without him I would not be the scientist or person that I am today.

I would also thank my committee members, Drs. Pete Mirabito, Rebecca Kellum and Michael Fried for their support in this process. I am also thankful to the faculty members of the Biology department who aided me surviving through this challenging journey. I am also very grateful to our wonderfully supportive office stuffs, Beverly Taulbee, Cheryl Edwards, Monica Decker, Seth Tylor and Melissa Justice.

I am very thankful to the past laboratory members of Rymond lab: Dr. Shatakshi Pandit, Dr. Li Zhang, Sudakshina and Mingxia for their valued assistance to cope up with the stress of graduate school in my initial days in this laboratory. Min Chen, a fantastic colleague and friend from this lab needs a special mention here. I must acknowledge the wonderful friends I have made in this journey: Dr. Li Hui, Sandeep, Dustin, Qian, Tom, Amber, Dave, Dan, Jason, Jiffin, Scott, Paul, Houfu, Wenhui, Ye, Xiaobo, Lakshmi, Wenwen, Steven, Lingfeng, Shreyas, Mansi, Drs. Travis Sexton, Tseten Yeshi, Deanna Morris, Peng Jiang, Amit Trivedi, Geng Wang and George Chaffins (late) for sharing the pain of this journey and making my life little less stressful.

Finally and most importantly my family needs a special acknowledgement: my father Rabindra Nath Banerjee, my mother Smita Banerjee, my aunt Pushi (Puspa Ghosh) and my wife Swagata Ghosh. I would like to thank my parents and aunt for coping up with life all these years without their only child. It is because of their tremendous support from 8000 miles away that I dared to dream this journey and now finally finishing it. The best thing that happened to me in this process is meeting my loving wife: Swagata, who is also my colleague in Rymond lab. A brilliant scientific mind in lab, a great friend and life partner at home who understands the pain of this journey, my life is beautiful every day because of her. I would also like to thank my parents-in-laws for their love and keeping faith in me.

Table of Contents

List of figures

List of Tables

CHAPTER 1: INTRODUCTION

1.1 Pre-messenger RNA splicing chemistry

Introns have to be removed from the primary RNA polymerase II transcripts to produce mature mRNAs, a biological process known as pre-mRNA splicing, as reviewed in (Chen and Cheng 2012). The pre-mRNA intron is defined by three consensus sequences largely conserved from yeast to mammals. In the yeast *Saccharomyces cerevisiae* the 5' splice site consensus is G/GUAPyGU (where / defines the exon-intron boundary), the branchpoint consensus is UACUAAC (where the underlined A is the branchpoint nucleotide) and the 3' splice site, PyAG/ (where / defines the intron-exon boundary). The splicing process occurs in the nucleus and takes place by two sequential trans-esterification reactions. First, the 2' hydroxyl group of the branchpoint adenosine acts as a nucleophile and attacks the 5'splice site. This cleavage reaction gives rise to two intermediary products: the free 5'exon and a lariat intermediate molecule comprised of the branched intron joined to the 3' exon. In the second cleavage reaction, the 3'hydroxyl group of the 5' exon attacks the phosphate at the 3' splice site and releases the intron and ligates the flanking exons. The released lariat intron is hydrolyzed at the 2'-5' phosphodiester linkage at the branch point (debranched) by the Dbr1 endonuclease and the linear intron subsequently degraded by endogenous nucleases (Chapman and Boeke 1991). The mature RNA is exported to the cytoplasm where it is translated into protein.

1.2 The spliceosome.

The pre-mRNA splicing reaction occurs within a complex macromolecular enzyme called the spliceosome. The yeast spliceosome is composed of 5 small nuclear RNAs (snRNAs) namely U1, U2, U4, U5 and U6 and approximately 100 proteins (Fabrizio, Dannenberg et al. 2009) and reviewed in (Will and Luhrmann 2011). Each snRNA is associated with a group of proteins and together they are called small nuclear ribonucleoproteins or snRNPs. Spliceosome assembly occurs in an ordered and stepwise manner that is conserved in most eukaryotes from yeast to mammals, reviewed in (Will and Luhrmann 2011, Chen and Cheng 2012). The assembly process initiates by the formation of a commitment complex (CC) when U1 snRNP binds to the 5' splice site in an ATP independent manner (Legrain, Seraphin et al. 1988). Following the formation of the commitment complex, the U2 snRNP binds to the branch point in an ATP dependent step. Next, the pre-assembled U4/U6-U5 tri-snRNP binds the pre-spliceosome. The U1 and U4 snRNPs are subsequently destabilized and released from the spliceosome. The removal of the U4 snRNP allows the U6 snRNA to base pair with the 5' splice site in place of the U1 snRNA. U6 snRNA also base pairs with U2 snRNA and this complex resides at the catalytic core of the spliceosome (Valadkhan 2005). A multi-subunit protein complex, the so-called NTC complex (Nineteen complex), stably binds the spliceosome and renders it catalytically active (Tsai, Fu et al. 2005). Interactions of the U5 snRNP with the 5' and 3' exons position the exons for joining in the second splicing reaction (Newman and Norman 1991, Newman and Norman 1992, Aronova, Bacikova et al. 2007, Crotti, Bacikova et al. 2007). After splicing occurs, the products are released

and the snRNPs are disassembled and recycled for subsequent rounds of splicing (Figure 1.1).

In addition to this canonical assembly pathway, a 45S protein complex containing U1, U2, U4, U5 and U6 snRNAs has been detected in yeast extracts that is splicing competent *in vitro* (Stevens, Ryan et al. 2002), suggesting that the spliceosome can function as a pre-assembled complex. To date, however, evidence for the existence of this "penta-snRNP" *in vivo* has not been found (Tardiff and Rosbash 2006).

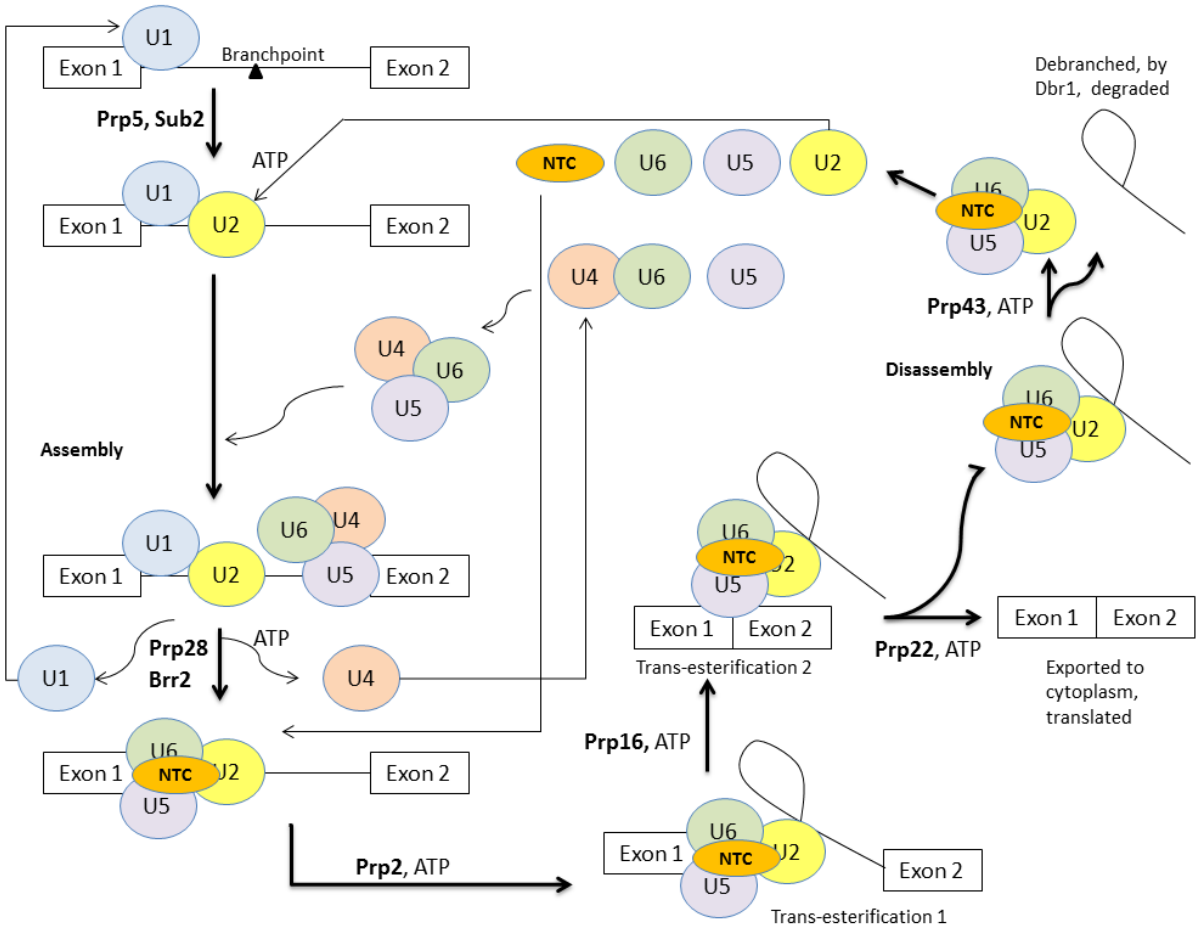

Figure 1.1. The spliceosome cycle. The spliceosome cycle is defined by a series of assembly, catalysis and the disassembly steps. In colored circles show the spliceosomal snRNP particles. The remodeling of spliceosome structure and composition are facilitated by a group of enzymes called the DExD/H box proteins depicted in bold.

1.3 The DExD/H box-helicases

The conserved steps of spliceosome assembly and subsequent catalytic activation of the spliceosome occur through sequential RNA-RNA and RNA-protein remodeling events. Many and perhaps all of remodeling events in the spliceosome are catalyzed by DExD/H proteins that function at distinct stages of the splicing pathway, reviewed in (Brow 2002, Rocak and Linder 2004, Cordin, Hahn et al. 2012) (Figure 1.1). DExD/H-box proteins utilize the energy derived from ATP binding and/or hydrolysis to promote the dissociation of RNA-RNA helices (i.e., helicase activity) or the dissociate protein factors from the pre-mRNA or associated snRNPs (i.e., RNPase activity). One example for remodeling catalyzed by a DExD/H box protein is the exchange of U1 snRNA with U6 snRNA at the 5' splice site of a pre-mRNA by Prp28 (Mathew, Hartmuth et al. 2008) and reviewed in (Maeder and Guthrie 2008).

The DExD/H-box family of proteins are defined by the presence of eight conserved protein motifs involved in RNA binding and hydrolysis, Figure 1.2 and reviewed in (Rocak and Linder 2004, Cordin, Banroques et al. 2006, Cordin, Hahn et al. 2012). There are eight known DExD/H proteins in yeast that assist the process of splicing, namely, Prp5, Sub2, Prp28, Brr2, Prp2, Prp16, Prp22 and Prp43 as reviewed in (Schwer 2001, Chen and Cheng 2012).

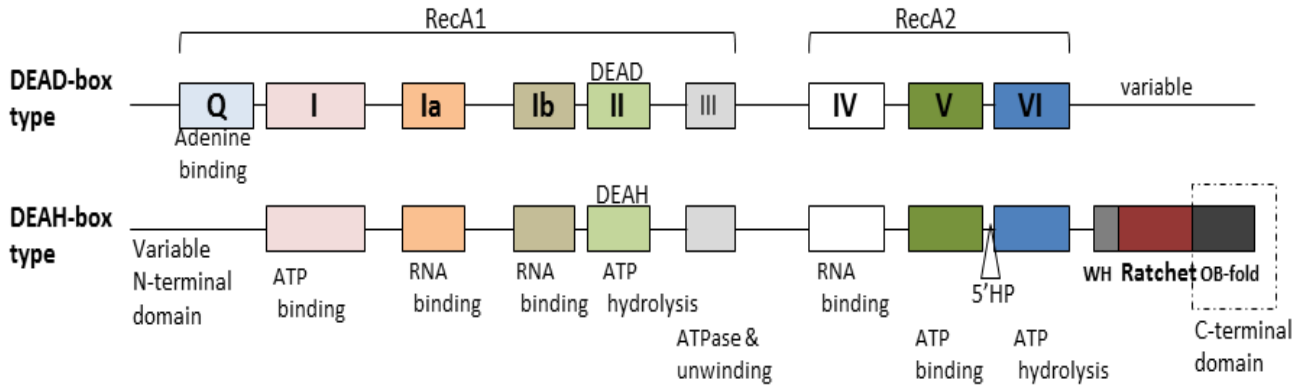

Figure 1.2. The conserved motifs of the DExD/H box family of proteins.

Shown here are the two major protein types present in the DExD/H family. The conserved motifs I-VI are represented by colored boxes with functionally important characteristics listed below. Motif I is also known as the Walker A motif, motifs Ia and Ib are part of domain I, motif II is also known as Walker B motif, motif III has been proposed to link ATPase and helicase activities and interacts with motif I and II. Motif IV, V and VI comprises the RecA2 domain. Conserved in both types are the 8 motifs, distributed over the RecA1 and RecA2 domains, the universally conserved helicase structures that contribute to ATP binding and hydrolysis. Unique to the DEAD-box type is the presence of the Q-motif at the N-terminal region, an adenine recognition motif. Unique to the DEAH-type is the presence of a hairpin-loop located between motif V and VI, and the presence of the winged-helix (WH) domain, ratchet domain, and the oligosaccharide/oligonucleotide binding domain at the CTD. The N-terminal domain in the DEAH-type proteins is variable.

In vivo, the activities of these DExD/H proteins are highly specific as mutations in each of the DExD/H proteins blocks spliceosome assembly at specific stages, reviewed in (Brow 2002, Will and Luhrmann 2011). *In vitro*, however, these proteins show little or no substrate specificity but display non-specific ATPase and helicase activities. For example Prp5, Prp28, Prp2, Brr2, Prp16, Prp22 and Prp43 each have been shown to have RNA dependent ATPase/NTPase activity (Schwer and Guthrie 1991, Kim, Smith et al. 1992, Strauss and Guthrie 1994, O'Day, Dalbadie-McFarland et al. 1996, Raghunathan and Guthrie 1998, Wagner, Jankowsky et al. 1998, Martin, Schneider et al. 2002) and Brr2, Prp16, Prp22, Prp43 also demonstrate ATP dependent RNA helicase activity (Wagner, Jankowsky et al. 1998, Wang, Wagner et al. 1998, Tanaka, Aronova et al. 2007, Maeder, Kutach et al. 2009). Clearly, such promiscuous enzymatic activity cannot happen in the cell as large amounts of ATP would be consumed by these enzymes without producing useful work or, even worse, impose ill-timed or cross-pathway disruptions in gene expression. The details of what regulates the temporal associations of the DExD/H proteins with the spliceosome and the catalytic activity of the DExD/H proteins are important open questions in the field.

Of the eight known DExD/H proteins in yeast, Sub2, Prp5, Prp28, Brr2 and Prp2 act prior to the first trans-esterification reaction in splicing. Sub2p has been proposed to dislodge BBP/Msl5 (branch-point binding protein) and its binding partner Mud2 which binds to the branch-point sequence, facilitating U2 snRNP association with the spliceosome (Kistler and Guthrie 2001, Wang, Zhang et al. 2008). Subsequently, Prp5 displaces Cus2 in an ATP dependent manner, further facilitating U2 snRNP binding to the spliceosome (Perriman, Barta et al. 2003). Prp28 and Brr2 are required for

spliceosome activation, by facilitating the release of the U1 and U4 snRNPs, respectively (Raghunathan and Guthrie 1998, Staley and Guthrie 1999) with Brr2 activity essential for unwinding U4/U6 intermolecular helices (Raghunathan and Guthrie 1998). Prp2 reorganizes the spliceosome to prompt changes in subunit association before the first trans-esterification reaction (Kim and Lin 1996) while Prp16 functions after the first trans-esterification reaction to promote the second step in splicing (Umen and Guthrie 1995, Schwer and Gross 1998, Schneider, Hotz et al. 2002). ATP dependent Prp22 helicase activity is required next for mature RNA release from the splicing complex (Schwer and Gross 1998, Campodonico and Schwer 2002, Schneider, Campodonico et al. 2004, Tanaka and Schwer 2005, Schwer 2008). Finally, Prp43 dissociates the post-catalytic spliceosome and releases the lariat intron (Arenas and Abelson 1997, Martin, Schneider et al. 2002).

Of the DExD/H-box splicing factors, Prp16, Prp5, Prp22 and Prp43 have been shown to have roles in splicing fidelity, that is, the selection of appropriate splice sites. An ATPase defective Prp16 mutant was found by the Guthrie group that suppresses mutations in the branchpoint sequence and accumulates lariat-intron molecules that form non-adenosine branchpoints showing that this protein is important for differentiating aberrant versus authentic branchsite selection (Burgess and Guthrie 1993). A role for Prp22 in proofreading exon ligation has been suggested from the observation that an ATPase defective Prp22 promotes the use of aberrant 3' splice site *in vitro* and *in vivo* (Mayas, Maita et al. 2006). Prp43 is the only DExD/H-box protein with spliceosome dissociation activity (Arenas and Abelson 1997). Diminished Prp43 activity suppresses growth defects in several spliceosome assembly mutants (Pandit, Lynn et al. 2006) and

allow for aberrant splice site selection on reporter genes (Mayas, Maita et al. 2010) suggesting that this protein is capable of dissociating defective spliceosome as well as the natural post-catalytic spliceosome (Pandit, Lynn et al. 2006) (and see below).

1.4 Prp43: a unique link between ribosomal RNA processing pathway and pre-messenger RNA processing pathway.

Ribosome biogenesis is the most highly regulated processes from bacteria to humans. Ribosomes are the most abundant proteins in the cell. In *S. cerevisiae*, there are approximately 150 rRNA genes, 137 Ribosomal Protein genes (RPG) and more than 200 additional non-ribosomal factors are involved in the process of ribosome production (Warner 1999, Henras, Soudet et al. 2008). Since this is a major energy consumer in the cell, depending on the environmental conditions, ribosomes are produced only when required. Thus, regulation of ribosome synthesis occurs at multiple levels: transcription of RPGs, transcription and processing of rRNAs, translation of ribosomal protein genes and turnover of rRNAs, ribosomal proteins and ribosomes (Leary and Huang 2001). Now, majority of intron containing RNA is ribosomal protein RNAs. The most abundantly spliced pre-mRNAs encode ribosomal proteins (RPs) (Ares, Grate et al. 1999).Under conditions in which translational resources are limiting; the spliceosome can rapidly down regulate the synthesis of new ribosomal components (Pleiss, Whitworth et al. 2007). Thus two huge parallel RNA processing pathways occur simultaneously in the nucleus: 1) ribosomal RNA processing, involving more than 200 protein factors and 75 small nucleolar ribonucleoproteins (snoRNPs) and 2) pre-messenger RNA processing involving 5 snRNAs and 80-100 proteins. The only common factors between these parallel RNA processing pathways are Prp43 and Snu13. Apart from its role in splicing

as described earlier, Prp43 is a critical component for maturation of both the small and large subunits of ribosomal RNA. The essential protein Snu13 binds to the U4 snRNA of the spliceosome and the box C/D snoRNAs of the pre-ribosomal processing pathway (Dobbyn, McEwan et al. 2007). I wanted to understand the biology in yeast how the activity of Prp43 is modulated in these two parallel RNA processing pathways.

1.5 Prp43 DEAH helicase structure

My research focuses on the regulation and function of the DExD/H box protein, Prp43. Independent protein crystallographic analyses from the Henry laboratory (Walbott, Mouffok et al. 2010) and from the Nielsen laboratory (He, Andersen et al. 2010) revealed that Prp43 has six conserved domains (Figure 4.2, Discussion): the N-terminal domain is unique to Prp43 and not observed in other members of this family, domains 2 and 3 (RecA1 and RecA2) which are the universally conserved helicase structures that contribute to ATP binding and hydrolysis, and Segments 4 and 5 defining the so-called winged-helix (WH) and ratchet domains respectively. The C-terminal domain (CTD) region of Prp43 includes an oligonucleotide/oligosaccharide-binding fold (OB fold) and this appears to be unique to the DEAH/RHA helicase subfamily of the DExD/H-box proteins. A model for domain contribution to in DExD/H box helicase function is provided in the Discussion.

1.6 Regulation of Prp43 in splicing.

The biological function of Prp43 in splicing is abetted by the 80 kDa Spp382 protein, also called Ntr1. *SPP382* was initially identified by our laboratory in a genetic screen for genes active in the discard pathway for non-productive spliceosomes (Pandit, Lynn et al. 2006) predicted by the kinetic proofreading model for spliceosome fidelity (Burgess and Guthrie 1993). Here, the *spp382-1* mutation was found as an extragenic suppressor of the spliceosome assembly mutant, *prp38-1* (Blanton, Srinivasan et al. 1992, Xie, Beickman et al. 1998). The official Saccharomyces Genome Database gene designation reflects its discovery as a Suppressor of *prp38-2*. Mutations of *SPP382* also suppress spliceosome assembly defects associated with mutations in the *PRP8* and *PRP19* splicing factors (Pandit, Lynn et al. 2006, Pandit 2009). Spp382 binds the splicing apparatus before Prp43, is required for Prp43 to join the spliceosome and stimulates the Prp43 helicase activity needed for spliceosome turnover (Tsai, Fu et al. 2005, Tanaka, Aronova et al. 2007).

The inhibition of Prp43 activity by *spp382* mutations is believed to suppress the *prp38-1* mutation by preventing the dissociation of a slow but active spliceosome (Pandit, Lynn et al. 2006). In other words, *spp382* mutations are believed to compensate for a kinetically impaired splicing reaction by stabilizing the partially active mutant splicing enzyme (that is, allowing the weakened enzyme longer to work before it is dissociated by Prp43). The specific protein or RNA features recognized by Prp43 to promote spliceosome turnover are unknown.

1.7 Regulation of Prp43 in ribosomal RNA processing.

Apart from its role in pre-mRNA splicing, Prp43 is also required for ribosomal RNA processing. Mutants of *PRP43* accumulate the 35S pre-rRNA precursor and show altered levels of several rRNA processing intermediates required for 18S and 25S -rRNA ribosome biogenesis (Lebaron, Froment et al. 2005, Combs, Nagel et al. 2006, Leeds, Small et al. 2006). Indeed, crosslinking experiments have identified Prp43 contacts on the 18S and 25S precursor molecules, demonstrating intimate association of Prp43 with this RNA (Bohnsack, Martin et al. 2009). Prp43 is not found in the mature ribosome, showing that its role is limited to the production of this critical cellular enzyme. The precise function of Prp43 in pre-rRNP biogenesis is unknown but given its intrinsic helicase activity, this protein might promote the dissociation of snoRNAs from pre-rRNAs, open up pre-rRNA features to permit snoRNA association or assist endonuclease cleavage or rRNA trimming by reorganization of the local RNP structure (Bohnsack, Martin et al. 2009, Pertschy, Schneider et al. 2009). As in pre-mRNA splicing, the details of pre-rRNP recognition by Prp43 remain unknown.

SQS1 (SQuelch of Splicing suppression) is a non-essential gene in yeast identified by Dr. Rymond's laboratory in a screen for genes that interact with *SPP382* (Pandit, Lynn et al. 2006). Although yeast cells lacking Sqs1 are viable and splicing competent, *SQS1* overexpression impairs pre-mRNA splicing and severely inhibits the growth of wild type yeast (Pandit 2009). Sqs1 interacts with both Prp43 and Spp382 in the yeast two hybrid assay (Pandit, Paul et al. 2009) and co-purifies with Prp43 (Gavin, Aloy et al. 2006). Prp43 and Sqs1 genetically interact with a pre-40S ribosomal protein Ltv1 and also with the endonuclease Nob1 that cleaves the 20S pre-rRNA to mature 18S rRNA

(Pertschy, Schneider et al. 2009). Sqs1 activates the Prp43 enzyme and appears to stimulate both the helicase and ATPase activity of Prp43 (Lebaron, Papin et al. 2009) unlike what was reported for Spp382 (Tanaka, Aronova et al. 2007).

Like Spp382 and Sqs1, a third pre-ribosomal particle protein, Pxr1, binds Prp43 in the two hybrid assay and can be purified from yeast in Prp43 complexes (Guglielmi and Werner 2002, Lebaron, Froment et al. 2005). Pxr1 is required for efficient rRNA biogenesis and is critical for normal growth of the yeast cells, as mutants bearing a *pxr1::KAN* null allele are severely growth impaired. Pxr1 is necessary for normal accumulation of Sqs1 *in vivo*, and for Sqs1 interaction with Prp43 and 20S rRNA (Lebaron, Papin et al. 2009). It remains unknown how Pxr1 mediates Prp43 biological function although Spp382 and Sqs1, and Pxr1 share a common structural feature likely relevant to this activity.

1.8 The G-patch domain: A connection between Prp43 activators and its dual function in splicing and rRNA biogenesis.

The data presented above provide evidence for a direct role of Prp43 in two fundamentally distinct RNA processing pathways, rRNA processing and pre-mRNA splicing. Consistent with this dual function, Prp43 appears to be at least 10 fold more abundant than the splicing-restricted DExD/H proteins like Prp2, Prp16 (Ghaemmaghami, Huh et al. 2003). Also, unlike the other splicing-associated DExD/H-box proteins, Prp43 localizes to nucleolus as well as to the nucleoplasm (Combs, Nagel et al. 2006). A direct physical association of Prp43 with the ribosome and the splicing machinery is supported by the co-purification of numerous splicing and ribosome

biogenesis co-factors with this DExD/H-box helicase (Lebaron, Froment et al. 2005, Gavin, Aloy et al. 2006).

A common feature of the Prp43 binding proteins, Spp382, Pxr1 and Sqs1 is that they all contain a single copy of the G-patch domain, an approximately 48 amino acid, glycine rich peptide found in select ribonucleoprotein (RNP) complexes (Aravind and Koonin 1999). *Saccharomyces cerevisiae* appears to have five proteins with the G-patch motif, namely, Spp382, Sqs1, Pxr1, Spp2 and Ylr271w. The location of the G-patch domain differs in these proteins; it is found in the N-terminus region of Spp382 (61-108 aa) and Pxr1 (25-72 aa) whereas the C-terminus of Sqs1 bears the G-patch domain (720-767 aa). In Spp2 it is located in the C-terminus (100-149 aa) and in YLR271w, it is found in the N-terminus (41-87 aa) segment. While three out of the five G-patch proteins bind Prp43, Spp2 binds Prp2 (Silverman, Maeda et al. 2004). We do not know anything about the interactions of the protein encoded by uncharacterized ORF *YLR271W*.

There is little structural insight into the G-patch domain and specific contribution of this motif to protein function remains obscure. The G-patch domain has been alternatively proposed as either a protein-protein or protein-RNA interaction module. Its nucleic acid binding property was demonstrated by experiments with the G-patch domain of the TgDRE protein by Frenal and colleagues (Frenal, Callebaut et al. 2006). Here this group used synthetic G-patch peptide and small RNA oligonucleotide to demonstrate G-patch domain-RNA binding by fluorescence anisotropy. The protein binding ability of the G-patch domain is suggested by studies with mutant G-patch protein derivatives which show reduced ability to bind purified recombinant Prp43 (Tanaka, Aronova et al. 2007, Lebaron, Papin et al. 2009).

1.9 The central hypothesis in my dissertation.

Correct partitioning of Prp43 between the rRNA and pre-mRNA processing pathway is important to optimize the biosynthetic potential of the cell. In addition, Prp43 activity, like that of other DExD/H-box helicases must be carefully regulated in these two pathways as pre-mature (or delayed) activation would presumably be detrimental. The biological basis for Prp43 partitioning between the splicing and rRNA processing pathways and the details of its temporal activation in each pathway Prp43 remain elusive. In this study, we seek to understand G-patch function in *S. cerevisiae*, through the characterization of the three G-patch proteins that bind Prp43, namely Spp382, Sqs1 and Pxr1. Based on studies from our laboratory and existing literature, I postulated a simple hypothesis that the Spp382/Sqs1/Pxr1 G-patches serve as alternate Prp43 binding surfaces that function with Prp43 in splicing or in rRNA processing (Figure 1.3). The Spp382, Pxr1 and Sqs1 proteins appear unrelated outside the G-patch region suggesting the possibility that their associations through these regions may be with different regions of Prp43.

The questions addressed in this study are:

1. Is the Pxr1 G-patch domain required for Prp43 association or activity?
2. Are the Spp382/Pxr1/Sqs1 G-patch peptides equivalent Prp43 binding surfaces?
3. Does G-patch sequence identify impact Prp43 function in splicing or rRNA processing?
4. Where is the G-patch binding domain within the Prp43 protein?

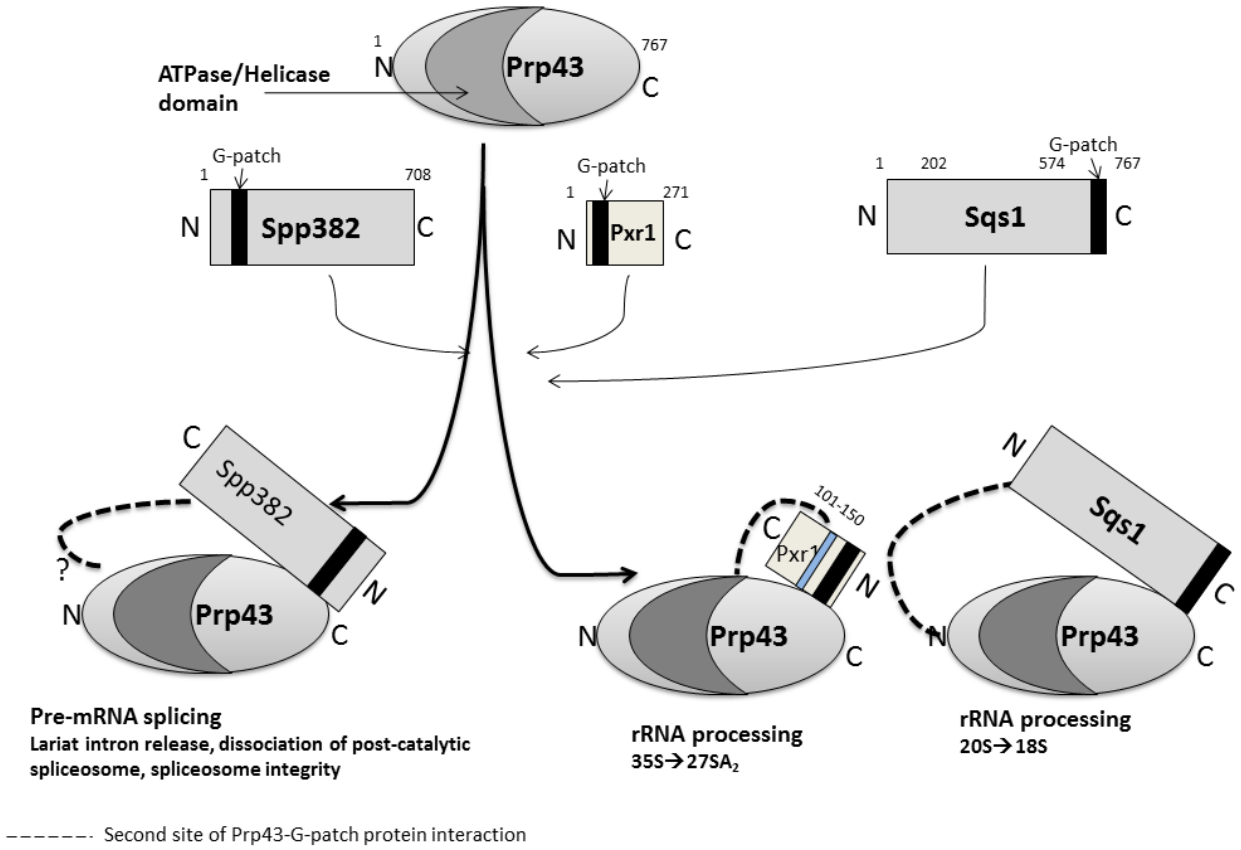

Figure 1.3. Model showing Spp382/Sqs1/Pxr1 G-patches serve as alternative Prp43 binding surfaces that function with Prp43 in splicing or in rRNA processing.

CHAPTER 2: MATERIALS AND METHODS

2.1 *Saccharomyces cerevisiae* methods:

2.1.1 Yeast strains.

Strain	Genotype	Reference
N19	*Mat A, ura3_0, trp1-289, leu2 _0, his3 _1, spp382*::*KAN*, and *p416-GAL1*::*spp382-4*	(Pandit, Lynn et al. 2006)
pJ69-4a	*Mat A, trp1-901 leu2-3,112, ura3-52, his3-200, gal4Δ, gal80 Δ , LYS2 : : GAL1–HIS3, GAL2–ADE2,met2 : : GAL7–lacZ*	(James, Halladay et al. 1996)
Δpxr1	*pxr1:: KAN, Mat A, leu2Δ0, ura3Δ0, his3Δ1*	Zhang and Rymond, unpublished.
BY4742	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0*	(Wang, Zhang et al. 2008)

2.1.2 Yeast transformation and maintenance.

Yeast strains were grown on rich media (1% bacto yeast extract, 2% bacto peptone, 2% sugar (glucose or galactose)) or on synthetic complete media with single or double amino acid dropouts at at 30C (F. Sherman, and et al. 1986). Yeast strains were transformed with 1-2 µg of plasmid DNA by standard techniques with lithium acetate treatment and heat-shock (Ito, Fukuda et al. 1983). The transformants were selected on dropout media based on the strain genotype and plasmid marker.

2.2 Plasmid shuffle assay

The characterization of gene mutations was done using the plasmid shuffle approach (Sikorski and Boeke 1991, Martin, Schneider et al. 2002) in which the indicated wildtype genes were cloned into the *URA3*-based vector and the mutant alleles cloned into the *LEU2* marked vector. Counter selection against the *URA3*-constructs was done on agar medium containing 0.75 mg/ml FOA (Martin, Schneider et al. 2002).

2.3 Colony growth assay

The relative growth of mutant and wildtype yeast strains was compared by colony growth assays performed on agar medium. To do this, the yeast strains were grown to saturation in a 1 ml culture. The cells were then collected by centrifugation and washed with sterile water once and finally resuspended in 1 ml of water. All cultures are normalized to OD_{600} 0.5 and four 10-fold serial dilutions prepared for each culture. 4 µl of each dilution was then spotted on indicated medium and incubated at the specified temperature and length of time. After incubation, the plates were scanned using an HP scanner (HP Scanjet G4050) and the image saved in Adobe Photoshop CS5 file format.

2.4 Yeast-2 hybrid assay

Interactions between proteins pairs were scored by the yeast-2 hybrid assay (Fields and Song 1989). The genes of interest were fused to the Gal4 activation domain (AD) and to the Gal4 DNA binding domain (BD) in the plasmids pACT2 (GeneBank ID U29899) and pAS2 (Harper, Adami et al. 1993), respectively. These constructs were then co-transformed in the yeast strain pJ69-4a (James, Halladay et al. 1996) harboring reporter genes *HIS3* under control of a *GAL1* promoter, *ADE3* gene under *GAL2* promoter and a LacZ gene under *GAL7* promoter. Interaction between Gal4 AD fused and Gal4 BD fused genes was measured as a function of yeast colony growth on a medium lacking histidine and containing the indicated amount of 3-amino triazole and/or by the β-galactosidase assay.

2.4.1 β-galactosidase assay

The pJ69-4a host harboring the Gal4 DNA binding domain and activation domain constructs were grown overnight in leucine, tryptophan dropout medium to OD_{600} ~ 1. 1 ml of culture was collected by centrifugation, washed once with 2 ml sterile water, resuspended in 1 ml Z buffer ($Na_2HPO_4.7H_2O$, 16. 1 g; $NaH_2PO4.H_2O$, 5.5 g; KCl, 0.75 g; $MgSO_4.7H_2O$, 0.246 g, β-mercaptoethanol, 2.7 ml; distilled water to make a final volume of 1 liter; adjusted the pH to 7, then stored at 4^{o}C until used) and 3 drops of chloroform were added. The mixture was vortexed for 10 seconds and incubated at room temperature for 5 minutes. The enzyme assay was started by adding 200 µl of Z buffer containing ortho-nitrophenyl- β-galactosidase substrate (4 mg/ml). When the yellow reaction product became obvious, the time was noted and the reaction was quenched by

addition of 500 µl of 1 M sodium bicarbonate (pH 8) and placing the sample on ice. The mixture was centrifuged for 10 minutes at 10,000 rpm to collect cellular debris.

The colorimetric measurement was done at OD_{420} and the β-galactosidase units were calculated as: 1000 X $O.D._{420}$ / $O.D._{600}$ of assayed culture X volume assayed (ml) X time in minutes.

2.5 Bacteria *Escherichia coli* methods:

2.5.1 *E. coli* strains used:

Strain	Genotype
DH5-α	supE44 Δ(lacU169 (Φ80d lacZ ΔhsdR17 recA1 endA1 gyrA96 thi-1 relA1(from Roche)
TG1	supE hsd Δ5 thi Δ(lac-proAB) F'[traD36proAB+ laclq lacZΔM15] (from Roche)

2.5.2 *E. coli* transformation and plasmid mini-preparation:

E. coli strains TG1 or DH5- α were grown in 2XYT medium (1.6% bacto tryptone, 1% bacto yeast extract, 0.5% NaCl, pH 7.5) or in Luria Broth (LB) medium (Sambrook J, Fritsch EF et al. 1989) at 37°C with ampicillin (100 mg/L) added when needed. *E. coli* transformation with plasmid DNA was done by making the cells competent with calcium chloride treatment followed by heat-shock at 42°C (Sambrook J, Fritsch EF et al. 1989). Small scale DNA plasmids were isolated from saturated cultures (1-2 ml) of *E. coli* by the alkaline lysis method (Sambrook J, Fritsch EF et al. 1989). The integrity and yield of plasmid recovery was tested by agarose gel electrophoresis. The plasmids for DNA

sequencing were isolated using Qiagen Plasmid Mini Kit (Catalogue #12125) following the vendor's protocol.

2.6 Plasmid construction:

2.6.1 Plasmids for functional reconstitution of *SPP382*:

Inverse polymerase chain reaction (PCR) was done using Ycplac111-*SPP382* template (Pandit, Lynn et al. 2006) with Phusion High-Fidelity DNA Polymerase (New England Biolabs Inc.) and primers having terminal SacII restriction sites flanking the G-patch domain (see Table 2.1). The PCR product was gel purified using a QIAquick Gel Extraction kit (Qiagen, Catalogue #28704). The purified DNA was digested with SacII, self-ligated and the transformed in *E. coli* to create Ycplac111-*spp382ΔG-patch.* To "add back" the missing domain, the Spp382, Pxr1 and Sqs1 G-patches were PCR amplified from yeast genomic DNA (strain BY4742), using gene-specific primers containing terminal SacII restriction sites. The PCR products were gel purified, digested with SacII, and ligated to the SacII linearized and dephosphorylated Ycplac111-*spp382ΔG* vector. Every mutagenesis (or cloned gene) construct was confirmed by PCR, restriction digestion and finally DNA sequence analysis (ACGT Inc. DNA Sequencing Services).

2.6.2 Construction of *PXR1* chimeras:

The wildtype *PXR1* gene was amplified from yeast genomic DNA along with 300 bps of upstream and 300 bps of downstream flanking sequence using gene specific primers (PXR1 natproKpn1F and PXR1 natproKpn1R) with terminal KpnI restriction sites on both ends. This fragment was ligated to the yeast centromeric plasmid Ycplac111 at

the vector KpnI site. This construct was confirmed by DNA sequence analysis and by complementation of the *pxr1::KAN* mutation (Zhang and Rymond, 2007, unpublished).

To create Ycplac111-*pxr1*ΔG-patch, the Ycplac111-*PXR1* plasmid was used for PCR reaction with primers (without any restriction site, Inv PXR1 delG For and Inv PXR1 delG Rev), flanking the G-patch domain. The PCR product was gel purified and phosphorylated with T4 polynucleotide kinase (PNK) (from NEB) and self-ligated. To create the chimeric constructs, linear unphosphorylated the Ycplac111-*pxr1*ΔG plasmid was ligated to the phosphorylated Spp382, Sqs1 or cognate Pxr1 G-patch domain prepared by PCR with primers (Pxr1 G-patch F and Pxr1 G-patch R; Spp382 G-For and Spp382 G-Rev; SQS1 G-For and SQS1 G-Rev).

2.6.3 Chimeric *SPP382* constructs for the Y2H assay.

The strategy for removing the G-patch and replacing with G-patches from Pxr1 and Sqs1 was identical to that described above for Ycplac111-s*pp382ΔG*. In this case, however, the starting template was pACT-*SPP382* (Pandit 2009).

2.6.4 *PXR1* deletion mutants for the Y2H assay.

The full length *PXR1* gene was first PCR amplified from yeast genomic DNA of strain BY4742 using *PXR1* SmaI F and *PXR1* SmaI R primers and then cloned in the SmaI site of the pACT2 vector as described earlier. The *PXR1* segments were removed from the pACT2-*PXR1* plasmid using inverse PCR mutagenesis with primers flanking the desired sites of deletion (see Table 2.1, primers 33-41).

2.6.5 *PRP43* deletion mutants for Y2H assay with G-patch proteins.

The *PRP43* deletion mutants were constructed by inverse PCR as described above. The starting template here being pAS2-*PRP43* (Pandit 2009) and the primers (Table 3, primers 50-65). The Prp43 isolated domain clones were made by PCR amplifying Prp43 NTD (78, 79), Ratchet (80, 81) and WH (82, 83) as NdeI fragment and cloned in the NdeI site of the pAS2 vector. The Prp43 RecA1 (66, 75), RecA2 (67, 68) and CTD (69, 70) were PCR amplified as BamH1 fragment and cloned in the BamH1 site of the pAS2 vector. Primer pairs listed in brackets, see Table 2.1 below.

2.6.6. pACT2-Isolated G-patches for the Y2H assay

The G-patch coding sequences were amplified from pACT-*SPP382*, pACT-*SQS1* (Pandit 2009) or pACT2-*PXR1* (this study) by PCR using primers *SPP382* G BamH1 end F and *SPP382* G BamH1 end R, *PXR1* G w BamH1 end F and *PXR1* G w BamH1 end R, *SQS1* G w BamH1 end F and *SQS1* G w BamH1 end R. These fragments were cloned in the BamH1 site of the pACT2 vector.

2.6.7 Site directed mutagenesis.

For site directed mutagenesis of the *PXR1* G-patch, the *PXR1* G-patch was first PCR amplified from genomic DNA using primers (*PXR1* Gw BamH1 end F and *PXR1* Gw BamH1 end R) and then blunt ended with Mung Bean Nuclease according to the vendors instructions (New England Biolabs Inc.). This PCR fragment was cloned into the SmaI site of pTZ18U (BIORAD) as a blunt-ended fragment. Site directed point mutations were introduced by inverse PCR with mutagenic oligos (see Table 3) using pTZ18U-

PXR1 G-patch as the template. This mutagenized DNA was then transferred to yeast centromeric plasmid Ycplac111 or yeast 2-hybrid plasmid pACT2 as a SacII fragment.

Table 2.1: **Primer list for plasmid construction.**

#1-18: *SPP382* cloning and mutagenesis:

	Name	Oligo sequence (5'>3')	Purpose
1	Inv SPP382 For	AAAAAA *CCGCGG* ACCAACTCCAGT AAC TAT CACA	PCR deletion of G-patch (61-108 aa) from SPP382 (introduce SacII site)
2	Inv SPP382 Rev	AAAAAA *CCGCGG* CTTCGTTAA TTTTGAG ATCG	PCR deletion of G-patch (61-108 aa) from SPP382 (introduce SacII site)
3	SPP382 G-patch For	AAAAAA *CCGCGG*ACATAT GGTATTGGTGCGAA	PCR amplification of SPP382 G-patch (61-108 aa) with SacII ends
4	SPP382 G-patch Rev	AAAAAA*CCGCGG*ATTTTGA AAACATTCCTAGACCT	PCR amplification of SPP382 G-patch (61-108 aa) with SacII ends
5	PXR1 G-patch For	AAAAAA *CCGCGG* ACC TCG AGA TTC GGG CACCA	PCR amplification of PXR1 G-patch (25-72 aa) with SacII ends
6	PXR1 G-patch Rev	AAAAAA*CCGCGG*TTTTAAT TTAGCACCGAG CCCAAC	PCR amplification of PXR1 G-patch (25-72 aa) with SacII ends
7	SQS1 F SacII	*CCGCGG* AACGAAAATATCGGTAGAA GA	PCR amplification of SQS1 G-patch (720-767 aa) with SacII ends
8	SQS1 G-patch Rev	AAAAAA *CCGCGG* ACTTTCACTGTGTCTTAAAC C	PCR amplification of SQS1 G-patch (720-767 aa) with SacII ends
9	Inv F PXR1-H55P	ATGAATTCGAACACTTCGC CTATC	Point mutagenesis of PXR1 G-patch
10	Inv R PXR1-H55P	GGGGGATAACCCCAGACCC AT	Point mutagenesis of PXR1 G-patch
11	InvF PXR1 R27G	TTT AAC CAC GGG CTT AGG GCT	Point mutagenesis of PXR1 G-patch
12	InvR PXR1 R27G	GAT CTT TTCAAA CCT ACC TTT GGG	Point mutagenesis of PXR1 G-patch

13	InvF PXR1 K57E	AAC ACT TCG CAT ATC GAA GTG	Point mutagenesis of PXR1 G-patch
14	InvR PXR1 K57E	CGA ATT CAT GGG GGA TAA CCC	Point mutagenesis of PXR1 G-patch
15	InvPxr1 D62M-F	ATGGACAACGTTGGGCTCG GTGCT	Point mutagenesis of PXR1 G-patch
16	InvPxr1 D62M-R	CTTAATTGACACTTTGATAT GCGA	Point mutagenesis of PXR1 G-patch
17	InvPxr1P 48G-F	GGCATGAATTCGAACACTT CGCAT	Point mutagenesis of PXR1 G-patch
18	InvPxr1P 48G-R	GGATAACCCCAGACCCATA CC	Point mutagenesis of PXR1 G-patch

#19-44: *PXR1* cloning and mutagenesis:

	Name	Oligo sequence (5'>3')	Purpose
19	PXR1 natproKp n1F	AAAAA *GGTACC* CTT AGCGTAAACAGTAACTGCG TGC	PCR amplification of PXR1 with natural promoter, KpnI ends (500bps up/125bps dn)
20	PXR1 natproKp n1R	AAAAA *GGTACC* TTTTTTCTTTGGCACCGGGG	PCR amplification of PXR1 with natural promoter, KpnI ends (500bps up/125bps dn)
21	PXR1 SmaI F	AAAAA *CCCGGG* G ATG GGT TTG GCA GCT ACA AG	PCR amplification of PXR1 with SmaI ends
22	PXR1 SmaI R	AAAAA *CCCGGG* CTA TTA GTC GTT TGT TAT CAT	PCR amplification of PXR1 with SmaI ends
23	Inv PXR1 For	AAAAAA *CCGCGG* GTCGTTACTCCATGCCGTAT TT	PCR deletion of G-patch (25-72 aa) from PXR1 (introduce SacII site)
24	Inv PXR1 Rev	AAAAAA *CCGCGG* CGT AAA GAC AAA AAA GACGAG	PCR deletion of G-patch from PXR1(25-72 aa) (introduce SacII site)
25	Inv PXR1 delG For	GTCGTTACTCCATGCCGTAT TTCT	PCR deletion of G-patch from PXR1 (no restriction site)
26	Inv PXR1 delG Rev	CGT AAA GAC AAA AAA GACGAGTTT	PCR deletion of G-patch from PXR1 (no restriction site)

27	Pxr1 G-patch F	ACCTCGAGATTCGGGCAC	PCR amplification of PXR1 G-patch w/o restriction site
28	Pxr1 G-patch R	TTTTAATTTAGCACCGAGC CCAAC	PCR amplification of PXR1 G-patch w/o restriction site
29	Spp382 G-For	ACATAT GGTATTGGTGCGAAG	PCR amplification of SPP382 G-patch w/o restriction site
30	SPP382 G- Rev	ATTTTGAAAACATTCCTAG ACC	PCR amplification of SPP382 G-patch w/o restriction site
31	SQS1 G-For	AAC GAAAATATCGGTAGAAGA	PCR amplification of SQS1 G-patch w/o restriction site
32	SQS1 G-Rev	ACTTTCACTGTGTCTTAAAC C	PCR amplification of SQS1 G-patch w/o restriction site
33	Invr Pxr1 del7-101	AAAAA*CCGCGG*TCTTGTAG CTGCCAAACCCAT	PCR deletion of PXR1 fragment (Δ7-101 aa) and introduce SacII site
34	Inv fPxr1del7 -101	AAAAA*CCGCGG*GAAAGCA AAATATCAGAGGAA	PCR deletion of PXR1 fragment (Δ7-101 aa) and introduce SacII site
35	Invr PXR1del 25-150	AAAAA*CCGCGG* GTCGTTACTCCATGCCGTAT T	PCR deletion of PXR1 fragment (Δ25-149 aa) and introduce SacII site
36	Invf PXR1del 25-150	AAAAA*CCGCGG*AAGAAAC GAAAGAGGGAA	PCR deletion of PXR1 fragment (Δ25-149 aa) and introduce SacII site
37	Invr PXR1 del101-226	AAAAA*CCGCGG*CTTTCCGT TCAGCCTACC	PCR deletion of PXR1 fragment (Δ102-226 aa) and introduce SacII site
38	InvF PXR1 del101-226	AAAAA*CCGCGG*TCTAGTGA ATCTGCATCC	PCR deletion of PXR1 fragment (Δ102-226 aa) and introduce SacII site
39	InvRPXR 1del150-226	AAAAA*CCGCGG*CGCGTTGG AGTAATTTCT	PCR deletion of PXR1 fragment (Δ150-226 aa) and introduce SacII site
40	InvrPXR 1del226-265	AAAAA*CCGCGG*TTTTAATT TTGACGTCTT	PCR deletion of PXR1 fragment (Δ227-265 aa) and introduce SacII site

41	InvfPXR 1del226-265	AAAAA*CCGCGG*TTTATGAT AACAAACGAC	PCR deletion of PXR1 fragment (Δ227-265 aa) and introduce SacII site
42	Pxr1_101 -150aa_F	AAAAA*GAATTC*GAGAAAGC AAAATATCAGAGGAATTG	PCR amplification of PXR1 (102-150 aa) with EcoR1 ends
43	Pxr1_101 -150aa_R	AAAAA*GAATTC*CTACTTCGC GTTGGAGTAATTTCTTAA	PCR amplification of PXR1 (102-150 aa) with EcoR1 ends
44	Pxr1_226 aa_R	AAAAA*GAATTC*CTATTTTAA TTTTGACGTCTTTTTCGT	PCR amplification of PXR1 (102- 226 aa) with EcoR1 ends

#45-68: *PRP43* cloning and mutagenesis:

	Name	Oligo sequence (5'>3')	Purpose
45	Inv Prp43del 7-94F	AAAAA*CCGCGG*GTACATGC CCAGAGAGATGAG	PCR deletion of PRP43 NTD (Δ7-94 aa) and introduce SacII site
46	InvPrp43 del 7-94R	AAAAA*CCGCGG*GAATCTTC TTTTGGAACCCAT	PCR deletion of PRP43 NTD (Δ7-94 aa) and introduce SacII site
47	InvPrp43 del 94-270F	AAAAA*CCGCGG*ACATATCC AGTTGAGCTATAC	PCR deletion of PRP43 RecA1 (Δ95-270 aa) and introduce SacII site
48	InvPrp43 del 94-270R	AAAAA*CCGCGG*TGGCAATT CTCTTCTAATTTT	PCR deletion of PRP43 RecA1 (Δ95-270 aa) and introduce SacII site
49	InvPrp43 del270-457F	AAAAA*CCGCGG*TTACGTTC CAATTTATCCTCC	PCR deletion of PRP43 RecA2 (Δ271-457 aa) and introduce SacII site
50	InvPrp43 del270-457R	AAAAA*CCGCGG*TCTGCCCG GAACGGCAAGTAA	PCR deletion of PRP43 RecA2 (Δ271-457 aa) and introduce SacII site
51	InvPrp43 delWHnewF	AAAAA*CCGCGG*CTAACTAT AGTGGCCATGCTATCG	PCR deletion of PRP43 WH (Δ445-542) and introduce SacII site
52	InvPrp43 delWHnewR	AAAAA*CCGCGG*TGCCTCTT CAGTGTATAATCTGAA	PCR deletion of PRP43 WH (Δ445-542) and introduce SacII site
53	InvPrp43 del521-635F	AAAAA*CCGCGG*ACTACAGA CTATGAAAGCCCT	PCR deletion of PRP43 Ratchet (Δ522-635) and introduce SacII site

54	InvPrp43 del521-635R	AAAAA*CCGCGG*GGGAAAC TGGGATGCCAATCT	PCR deletion of PRP43 Ratchet (Δ522-635) and introduce SacII site
55	InvPrp43 delCTDnewR	AAAAA*CCGCGG*GATGTTGT CAAAGTATTTAGGGCT	PCR deletion of PRP43 CTD (Δ649-748 aa) and introduce SacII site
56	InvPrp43 delCTDnewF	AAAAA*CCGCGG*AGATTAAA CGAGTTGAAACAAGGT	PCR deletion of PRP43 CTD (Δ649-748 aa) and introduce SacII site
57	Prp43_NTD_F	AAAAA*CATATG*ATGGGTTC CAAAAGAAGATTC	PCR amplification of Prp43 NTD (1- 94 aa)with NdeI ends
58	Prp43_NTD_R	AAAAA*CATATG*CTATGGCA ATTCTCTTCTAATTTTCAG	PCR amplification of Prp43 NTD (1- 94 aa) with NdeI ends
59	Prp43RecA1_F	AAAAA*GGATCC*GTGTACAT GCCCAGAGAGATGAGTTT	PCR amplification of Prp43 RecA1 (95- 270 aa) with BamH1 ends
60	Prp43_RecA1_R	AAAAA*GGATCC*CTATCTGC CCGGAACGGCAAGTAA	PCR amplification of Prp43 RecA1 (95- 270 aa) with BamH1 ends
61	Prp43RecA2_F	AAAAA*GGATCC*GTACATAT CCAGTTGAGCTATACTAT	PCR amplification of Prp43 RecA2 (271- 457 aa) with BamH1 ends
62	Prp43RecA2_R	AAAAA*GGATCC*CTAAATTT CTGGGTAACTTTGCTCTAT	PCR amplification of Prp43 RecA2 (271- 457 aa) with BamH1 ends
63	Prp43_WH_F	AAAAA*CATATG*TTACGTTCC AATTTATCCTCC	PCR amplification of Prp43 WH (458- 521 aa) with NdeI ends
64	Prp43_WH_R	AAAAA*CATATG*CTAGGGAA ACTGGGATGCCAATCT	PCR amplification of Prp43 WH (458- 521 aa) with NdeI ends
65	Prp43_Ratchet_F	AAAAA*CATATG*TTGGATCCT ATGCTAGCGGTG	PCR amplification of Prp43 Ratchet (522- 635 aa) with NdeI ends
66	Prp43_Ratchet_R	AAAAA*CATATG*CTAGTTTAA TTCTAGGTTGTAACG	PCR amplification of Prp43 Ratchet (522- 635 aa) with NdeI ends
67	Prp43CTD_F	AAAAA*GGATCC*GTACTACA GACTATGAAAGCCCTAAA	PCR amplification of Prp43 CTD (636-767 aa) with BamH1 ends
68	Prp43CTD_R	AAAAA*GGATCC*CTATTTCTT GGAGTGCTTACTCTT	PCR amplification of Prp43 CTD (636-767 aa) with BamH1 ends

#69-74: Cloning of G-patch domain:

	Name	Oligo sequence (5'>3')	Purpose
69	SPP382 BamH1 end F	AAAAA *GGATCC* GA ACATATGGTATTGGTGCGA AG	PCR amplification of SPP382 G-patch with BamH1 ends
70	SPP382 G BamH1 end R	AAAAA *GGATCC* TTA ATTTGAAAACATTCCTAGA CCT	PCR amplification of SPP382 G-patch with BamH1 ends
71	PXR1 Gw BamH1 end F	AAAAA *GGATCC* GA ACCTCGAGATTCGGGCAC	PCR amplification of PXR1 G-patch with BamH1 ends
72	PXR1 G BamH1 end-R	AAAAA *GGATCC* TTA TTTTAATTTAGCACCGAGC CCAAC	PCR amplification of PXR1 G-patch with BamH1 ends
73	SQS1 G w BamH1 end F	AAAAA *GGATCC* GA AACGAAAATATCGGTAGAA GA	PCR amplification of SQS1 G-patch with BamH1 ends
74	SQS1 G w BamH1 end R	AAAAA *GGATCC* TTA ACTTTCACTGTGTCTTAAAC C	PCR amplification of SQS1 G-patch with BamH1 ends

Note: Italicized nucleotide sequence represents restriction site.

Location of primers: Banerjee Primer Box-1 and Banerjee Primer Box-2.

2.7 RNA methods:

2.7.1 Northern blot analysis:

The yeast cultures were grown to $O.D._{600}$ approximately 0.4 in the indicated medium in 3-5 ml of culture volume. The cells were collected by centrifugation, washed twice with ice cold RE buffer (100 mM LiCl, 100 mM Tris-HCl pH 7.5, 1 mM EDTA) and the pellets stored at -80°C until needed. Total RNA was isolated by resuspending the cell pellet in 400 µl RE buffer, $2/3^{rd}$ volume of sterile glass beads (0.5mm diameter, BioSpec Products) and 400 µl phenol/chloroform/isoamyl alcohol (50:49:1) (PCI). The cells were broken in the Mini-Beadbreaker (BioSpec Products) using a chilled tube adaptor for 4 minutes. The cellular debris, glass beads and PCI were removed by centrifugation at 18,000 rcf for 3 mins at 4°C. The supernatant was collected and extracted for 3 times with PCI followed by chloroform extraction once. $1/5^{th}$ volume of 3 M sodium acetate, pH 5.2 was added followed by 2-3 volumes of 100% ethanol. The mixture was placed on dry ice for 15 minutes and then RNA collected by centrifugation at 14,000 rcf for 10 minutes. This RNA pellet was washed twice with 70% ethanol, dried and resuspended in 50 µl of RNase free water. This RNA was quantified by spectrophotometric analysis ($O.D._{260}$).

20 µg of RNA was resolved on a 1% agarose/formaldehyde gel (F. Sherman, and et al. 1986). A Random Primer Labeling Kit (Invitrogen) was used to make uniformly radiolabeled ^{32}P probes against the ribosomal protein *RPS17A* gene or *ADE3* gene, following the manufacture's protocol. The radiolabelled probes were hybridized and washed under standard conditions (Sambrook J, Fritsch EF et al. 1989). The RNA bands were visualized with Typhoon 8600 Phosphoimager (GE Biosystems). Splicing

efficiency was accessed by determining the messenger RNA to precursor RNA ratio with ImageQuant software (Molecular Dynamics).

5’ end-labeled oligonucleotide probes were used to detect the 35S rRNA precursor (004), 23S and 20S intermediates (004), 27SA2 intermediate (003) and 25S (007), 18S (008), 7S (020) and 5.8S (017) rRNA (see Table 2.2). Overnight hybridization was done with the oligonucleotide probes in 6× SSC, 5 × Denhardt’s solution, and 1% SDS at 43°. The blots were washed in 1× SSC, 0.5% SDS at 43° (rRNA 003,007, 008, 017), 36° (rRNA 004) or 50°C (rRNA020) as previously described for these oligonucleotide probes (Kos and Tollervey 2005).

2.7.2 Transcriptional repression of functional Spp382 in vivo:

To deplete functional Spp382 in vivo, yeast strains bearing *p416-GAL1::spp382-4* was grown at 30°C in minimal media containing galactose to $O.D._{600}$ 0.2. The cells were harvested by centrifugation and washed with sterile water for three times and then replaced by a minimal media containing glucose and grown for 18hrs at 30°C. Cells were harvested at $O.D._{600}$ ~0.4 and RNA was isolated as described before.

Table 2.2: **DNA oligonucleotides used for rRNA processing pathway** (Kos and Tollervey 2005).

Oligo name	5'>3' sequence	rRNA identified
003	TGTTACCTCTGGGCCC	27SA2
004	CGGTTTTAATTGTCCTA	35S rRNA precursor, 23S and 20S
007	CTCCGCTTATTGATATGC	25S
008	CATGGCTTAATCTTTGAGAC	18S
017	GCGTTCTTCATCGATGC	5.8S
020	TGAGAAGGAAATGACGCT	27SA, 27SB, 7S

2.8 Bioinformatics analysis:

The putative G-patch protein homologs were identified by BLASTP (BLOSUM 62 matrix) at the National Center for Biotechnology Information (NCBI) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) using the full-length *S.cerevisiae* G-patch proteins as query sequences (sequence obtained from *Saccharomyces Genome Database*, http://www.yeastgenome.org/) against the non-redundant(nr) protein databases of indicated species. The sequence match with lowest E-value was selected as the most likely homolog. The putative homologous genes of the three G-patch proteins from 11 different species (*Saccharomyces cerevisiae, Schizosaccharomyces pombe, Arabidopsis thaliana, Drosophila melanogaster, Caenorhabditis elegans, Danio rerio, Xenopus laevis, Gallus gallus, Mus musculus, Bos Taurus* and *Homo sapiens*) were retrieved from the NCBI Protein Database (http://www.ncbi.nlm.nih.gov/protein/) using the accession number provided in Table 2, below. In some cases, the retrieved protein sequences included database annotations indicating the homology to the yeast protein (see Table 2.3). The G-patch coordinates for each protein were taken from NCBI protein database annotations. Multiple Sequence Alignment was performed with the G-patches of Spp382 (Sp), Pxr1 (Px) and Sqs1 (Sq) proteins across the 11 species using MUSCLE (Edgar 2004, Edgar 2004) (http://www.ebi.ac.uk/Tools/msa/muscle/)

Table 2.3 **NCBI protein accession number for G-patch proteins across 11 species.**

	Species	NCBI ID Spp382	NCBI ID Pxr1	NCBI ID Sqs1
1	*Saccharomyces cerevisiae*	AAB67507.1*	AAS56710.1α	NP_014175.1
2	*Schizosaccharomyces pombe*	NP_594091.1* Query coverage: 16%, E value: 3e-08, Identities = 24/56 (43%)	Q9URX9.1 α Query coverage: 43%, E value: 1e-24 Identities = 47/117 (40%)	NP_595527.1 β Query coverage :26%, E value: 4e-21, Identities = **66/212 (31%)**
3	*Arabidopsis thaliana*	NP_173150.1* Query coverage:6%, E value: 2e-08, Identities = 26/44 (59%)	NP_566957.1 α Query coverage:18%, E value: 5e-06, Identities = **22/51(43%)**	NP_850884.1 β Query coverage:14%, E value: 8e-10, Identities **=41/114(36%)**
4	*Drosophila melanogaster*	AAF52282.1* Query coverage:48%, E value: 8e-11, Identities = 82/347(23%)	NP_611495.1 α Query coverage:37%, E value: 3e-13 35/101 (35%)	NP_725443.1 β Query coverage:14%, E value: 8e-10 Identities = **41/114(36%)**
5	*Caenorhabditis elegans*	NP_496226.2* Query coverage:34%, E value: 7e-07, Identities = 65/268(24%)	NP_495955.1Δ Query coverage:45%, E value: 9e-14 Identities = **51/159(32%)**	NP_491200.1 β Query coverage:16%, E value: 4e-09, Identities =**47/134(35%)**
6	*Danio rerio*	NP_001002721.1* Query coverage:23%, E value: 8e-11, Identities = 61/208(29%)	AAH54670.1 α Query coverage:46%, E value: 1e-22, Identities = **54/141(38%)**	AAI65705.1 β Query coverage:15%, E value: 3e-13, Identities = **50/127(39%)**

7	*Xenopus laevis*	NP_001087642.1* Query coverage:69%, E value: 4e-12, Identities = 156/646(24%)	NP_001086366.1 α Query coverage:48%, E value: 8e-22, Identities = **54/132(41%)**	NP_001086761.1 β Query coverage:23%, E value: 8e-10 Identities = **68/215(31%)**
8	*Gallus gallus*	NP_001025836.1* Query coverage:27%, E value: 3e-11 Identities = 56/144(39%)	XP_420036.3Δ Query coverage:35%, E value: 4e-23 Identities = 42/97 (43%)	NP_001026217.1 β Query coverage:19%, E value: 4e-10 Identities =
9	*Mus musculus*	NP_061253.2* Query coverage:27%, E value: 6e-11, Identities = 83/286(29%)	NP_082504.1 α Query coverage:47%, E value: 1e-21, Identities = 57/128(44%)	EDL13048.1 β Query coverage:10%, E value: 1e-10 Identities = **34/77(44%)**
10	*Bos Taurus*	NP_001039495.2* Query coverage:40%, E value: 5e-11 Identities = 101/331(30%)	ABF57427.1 α Query coverage:35%, E value: 2e-23, Identities = 44/97 (45%)	NP_001193622.1 β Query coverage:18%, E value: 2e-09 Identities = **60/142(42%)**
11	*Homo sapiens*	NP_001008697.1* Query coverage:28%, E value: 5e-11, Identities = 88/299(29%)	AAS19507.1 α Query coverage:35%, E value: 7e-23, Identities = 43/97 (44%)	NP_060510.1 β Query coverage:9%, E value: 6e-11, Identities = **37/73(50%)**

Proteins listed above as "*" or "α" -genes are annotated at NCB1 as putative homologs of the yeast proteins (http://www.ncbi.nlm.nih.gov/homologene/); proteins listed as "Δ" and "β" means the proteins were not annotated as homologs in the database but were chosen based on the best BlastP score with the *S. cerevisiae* peptide.

2.9 Cloning and purification of a chimeric Spp382 (1-121 aa) peptide.

The Spp382 N-terminal 121 aa peptide containing the G-patch domain is sufficient to bind and activate Prp43 (Tanaka, Aronova et al. 2007). We cloned the DNA encoding this 121 aa segment into the pTBX1 (Catalog #E6901S) as a NdeI and SapI PCR fragment (B. Rymond, unpublished). This construct produces a protein that fuses a C-terminal self-cleavable intein tag followed by a chitin binding domain (CBD) tag joined to the 121 aa Spp382 peptide. The Spp382 G-patch domain was removed by inverse PCR (with primers having SacII site), and Pxr1, Pxr1 (H55P) or Sqs1 G-patch was added back in place of the cognate Spp382 G-patch using SacII restriction site.

Protein Purification: The wildtype or the chimeric Spp382 constructs were transformed in the *E. coli* strain ER2566 (Impact system manual, E6901, NEB). The *E. coli* culture was grown to $O.D._{600}$ ~0.5 at 37°C and protein expression was induced with 0.4 mM IPTG, and the culture was grown at 16°C for 18 hours. The cells were harvested by centrifugation and cell pellets were stored at -80°C. All subsequent steps were done at 4°C. Cell pellets from 1 liter of culture were resuspended in 100 ml of ice-cold of Cell Lysis Buffer (20 mM Tris-HCl, pH 8 .5; 500 mM NaCl; 1 mM EDTA; 0 .1% Triton X-100) and lysed by sonication. The cell debris was removed by centrifugation at 15,000 g for 30 minutes in a Sorval RC-5B rotor. The cleared lysate was mixed with 5 ml of a 50% chitin beads suspension equilibrated with column buffer (20 mM Tris-HCl, pH 8.5; 500 mM NaCl; 1 mM EDTA) for 30 mins. The chitin beads were then collected by centrifugation (4000 g, 3 mins); washed three times with three volumes of column buffer and loaded onto a 20 ml disposable column. The peptide-bound chitin beads were washed

with an additional 20 bed volumes of column buffer (adjusted to 1M NaCl) and then with 5ml of 50 mM NaCl containing column buffer for three times.

2.9.1 Binding assay:

The Prp43 expression construct was a kind gift of Dr. Beate Schwer and this protein was purified as previously described (Tanaka and Schwer, 2006). Prp43 (10 µg approximately) was rotated with peptide bound chitin beads (30 µg approximately) (1:6 molar ratio) for 1 hour in 4°C in final volume 100 µl in binding buffer (50 mM NaCl, 50 mM imidazole, 50 mM Tris-HCl at pH 7.4, 10% glycerol). Next, the chitin beads were collected by centrifugation, the unbound Prp43 was removed, the beads washed 3 times in 100 µl wash buffer (50-200 mM NaCl as indicated, 50 mM imidazole, 50 mM Tris-HCl at pH 7.4, 10% glycerol) and finally resuspended in 50 µl protein sample buffer (100 mM Tris-HCl, pH 6.8; 20% glycerol; 5% SDS; 5% β-mercaptoethanol 0.02% bromophenol blue). The samples were boiled for 10 mins to release the peptide along with the peptide-bound Prp43 prior to gel electrophoresis on a10% SDS-PAGE. The protein bands were visualized by staining the gel with Coomassie Blue followed by de-staining with a 50% methanol, 5% acetic acid solution. The de-stained gels were photographed with Fujifilm camera.

CHAPTER 3: RESULTS

3.1 The isolated G-patch from Spp382, Pxr1 and Sqs1 each interacts with full length Prp43 with the apparent activity Spp382>Sqs1>Pxr1.

The approximately 48 amino acid, glycine rich G-patch domain has been proposed as protein-protein or protein-RNA interaction module for RNP proteins (Aravind and Koonin 1999). It has been shown that in *Saccharomyces cerevisiae*, Spp382, Sqs1 and Pxr1 each interacts with Prp43 by genetic and biochemical experiments (Guglielmi and Werner 2002, Lebaron, Froment et al. 2005, Tsai, Fu et al. 2005, Boon, Auchynnikava et al. 2006, Tanaka, Aronova et al. 2007, Lebaron, Papin et al. 2009, Pandit, Paul et al. 2009). Here we used the yeast 2 hybrid (Y2H) assay to test the hypothesis that the G-patch domain is sufficient for Prp43 interaction. The G-patch domain boundaries were selected based on previous literature (Tanaka, Aronova et al. 2007, Lebaron, Papin et al. 2009, Pandit 2009). Each of the isolated G-patch domains of Spp382, Pxr1 or Sqs1 was fused to the Gal4 activation domain (AD) and full length Prp43 was fused to the Gal4 DNA binding domain (BD) in the plasmids pACT2 (GeneBank ID U29899) and pAS2 (Harper, Adami et al. 1993), respectively. These constructs were then co-transformed in the yeast strain pJ69-4a (James, Halladay et al. 1996) harboring reporter genes *HIS3* under the control of a *GAL1* promoter, *ADE3* gene expressed from the *GAL2* promoter and a LacZ gene expressed from the *GAL7* promoter. Interaction between the G-patch domain and Prp43 was scored as yeast colony growth on a medium lacking histidine.

Figure 3.1 shows the result of my yeast-2 hybrid (Y2H) study between *PRP43* and the full-length G-patch proteins or the isolated individual G-patch domains. The rows show the growth cultures spotted on the indicated medium in 10-fold serial dilutions,

starting with an initial adjusted value of OD_{600} of 0.5. The left-hand panel shows simple growth, selecting for the plasmid markers but not reporter gene trans-activation. The right-hand panel shows yeast cell growth on the medium requiring successful Y2H trans-activation. The pattern of growth here suggests a gradation of Prp43-G-patch protein interaction of Spp382 >Sqs1>Pxr1. The isolated G-patches show a similar pattern although with reduced overall activity with the Prp43-Pxr1 G-patch interaction being modestly above the background observed with the empty vector, pACT2. It is also possible that this difference is yeast cell growth might also reflect differences in protein concentration or different levels of interfering interaction.

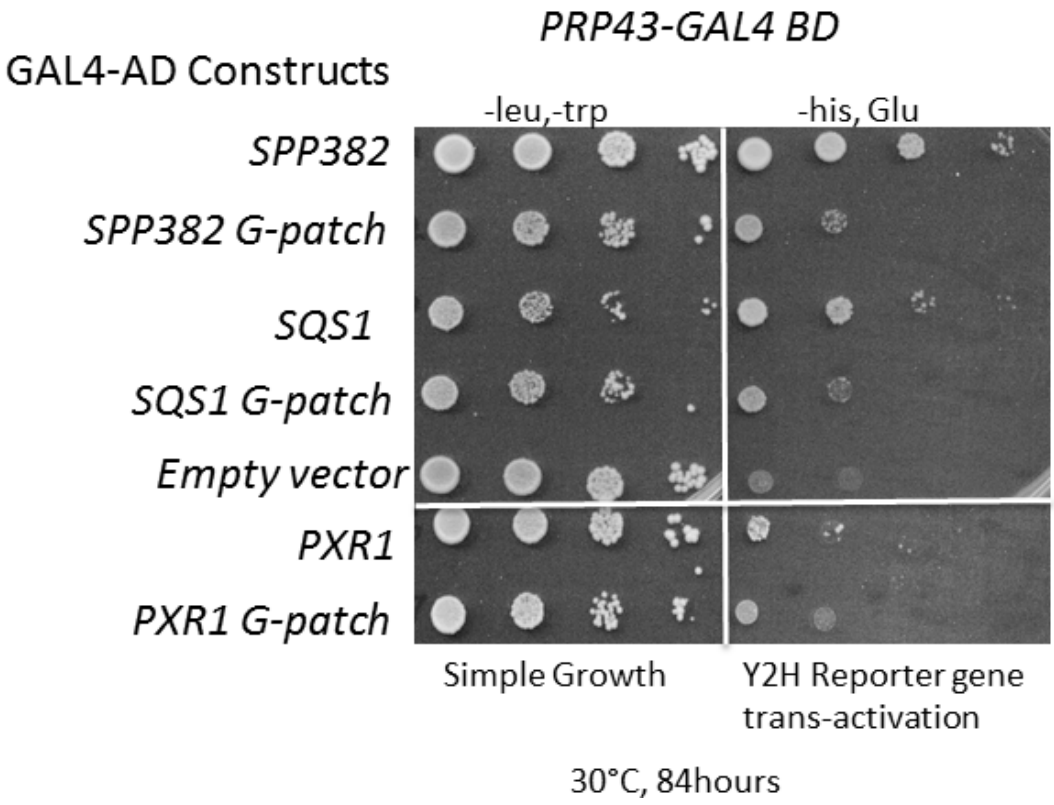

Figure 3.1. The G-patch is sufficient for modest but detectable Prp43 interaction in yeast 2-hybrid assay. Each row shows the indicated pJ69-4a yeast transformants spotted on indicated medium in 10 fold serial dilutions. The left panel shows growth on media to select for Y2H plasmids; right panel shows growth on medium lacking histidine where the GAL1-*HIS3* reporter gene trans-activation is required. Growth of yeast colonies bearing full length Prp43 and isolated G-patches is compared to the empty vector negative control.

3.2 The G-patch is critical for DExD/H –box protein Prp43 interaction with Spp382 but not for its interaction with Sqs1 or Pxr1.

The G-patch domain was suggested to be essential for Spp382 to interact with Prp43 based on prior genetic and biochemical experiments (Tsai, Fu et al. 2005, Boon, Auchynnikava et al. 2006, Tanaka, Aronova et al. 2007). The Sqs1 G-patch domain, in contrast, is not critical for Prp43 interaction (Lebaron, Papin et al. 2009, Pandit, Paul et al. 2009) and a second tethering surface is located within the N-terminal 201 amino acids of Sqs1 (Lebaron, Papin et al. 2009). For the third G-patch protein binding partner of Prp43, Pxr1, it is unknown whether the G-patch is needed or not for Prp43 interaction. Pxr1 co-purifies with Prp43 in a yeast complex containing both ribosome biogenesis and splicing factors and interacts in a yeast-2 hybrid assay (Lebaron, Froment et al. 2005). We used yeast 2-hybrid assay to test the hypothesis that Pxr1 G-patch domain is required for Prp43 interaction.

The Pxr1 G-patch was deleted from this gene by inverse PCR and the resultant pGAL4-AD -*pxr1ΔG* construct was transformed in the yeast strain pJ69-4a (James, Halladay et al. 1996), harboring full length GAL4-BD-*PRP43*. Likewise, G-patch deletion derivatives were created for the Spp382 and Sqs1 G-patch proteins and these were assayed for Y2H interaction with full length Prp43. Figure 3.2 presents the results of this study. The right panel shows that the Spp382 G-patch is critical for Prp43 interaction as the *spp382ΔG* grows no better than the empty vector control, corroborating earlier peptide binding and mutagenesis studies (Tsai, Fu et al. 2005, Tanaka, Aronova et al. 2007). The Sqs1 G-patch is not critical for Prp43 interaction, also supporting earlier findings (Lebaron, Papin et al. 2009, Pandit, Paul et al. 2009). Removal of the Pxr1 G-

patch was found not to detectably impair its interaction with Prp43, indicating that a second binding site for Prp43 exists within Pxr1. We note a somewhat stronger Prp43 Y2H response with the *pxr1ΔG* construct than with the unmutagenized *PXR1*. While other explanations are possible, like increased protein expression with the *pxr1ΔG* construct, this might reflect an improved Prp43 interaction due to unmasking of this secondary site for Prp43 interaction within Pxr1 upon G-patch removal.

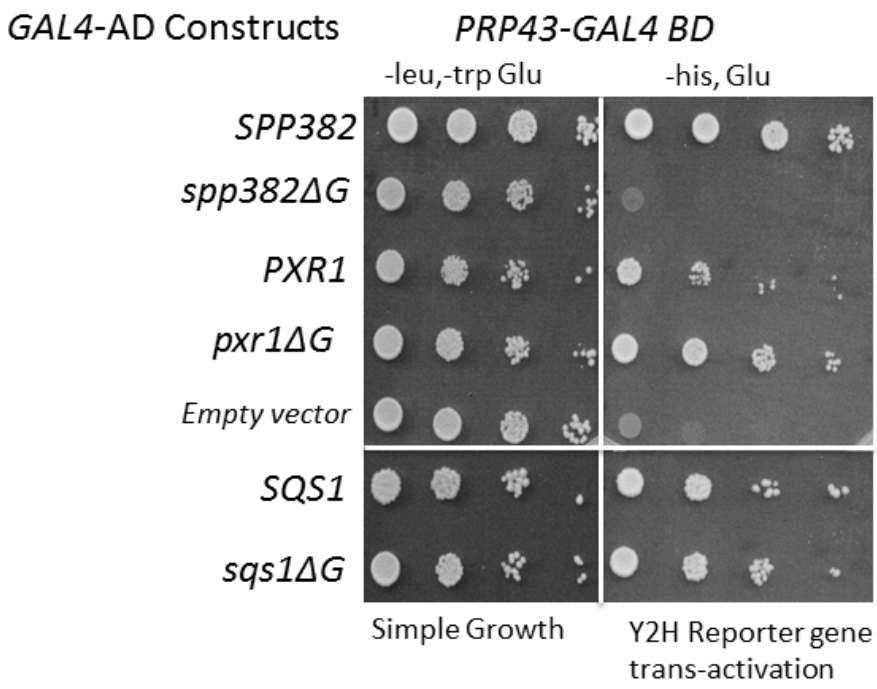

Figure 3.2. The G-patch is critical for Prp43 Y2H interaction with Spp382 but not for Prp43 interaction with the Sqs1 or Pxr1 proteins. Each column shows 10 fold serial dilutions on indicated medium for simple growth (left) or *GAL1-HIS3* reporter gene transactivation (right). The plates were incubated at 30°C for 84 hours.

3.3 Pxr1 segment 102-149 is a second site of association with Prp43.

The Pxr1 G-patch is not required for Prp43 interaction in a yeast-2 hybrid assay, opening up the question where the other Prp43 binding site(s) is located within Pxr1. To address this, I created a series of *pxr1* deletion derivatives by PCR (Figure 3.3.1) in the yeast 2-hybrid construct pACT2-*PXR1* (Gal4 activation domain fused Pxr1) and then scored each for interaction with Prp43 by the Y2H assay.

The Pxr1 regions deleted for this Y2H assay are illustrated in Figure 3.3.1 (indicated by Δ) and the Y2H results with the full length Prp43 protein presented in Figure 3.3.2. Deletion of the Pxr1 N-terminal domain (Δ7-101) has little or no detectable impact on Prp43 interaction compared to what is seen with the full length Pxr1 protein. Likewise, removal of the KKE/D region (150-226 aa) (Guglielmi and Werner 2002) or the C-terminal domain (CTD), i.e., residues 150 to 265 aa (Δ 150-265) has minimal impact on Prp43 interaction, showing that the region 102-149 aa of Pxr1 may sufficient to interact with Prp43. Removal of multiple other peptide fragments that include residues 102-149 aa (Δ102-226, Δ102-149 and Δ7-149) greatly impairs Prp43 interaction suggesting that the 48 amino acids in common that lie between residues 102 to 149 might harbor a second Prp43 binding site. To learn if the Pxr1 102-149 segment is sufficient for interaction, we fused this segment to the Gal4 AD and scored its Y2H interaction with Prp43. Figure 3.3.3 shows that this small 48 amino acid segment (i.e., Pxr1 102-149), fused to Gal4 AD, interacts with Prp43, although with reduced efficiency compared to full length Pxr1. A larger segment of Pxr1 (102-226 aa) interacts with Prp43 somewhat better suggesting that the segment from amino acids 150 to 226 might contribute this protein's stability or with its interaction with Prp43.

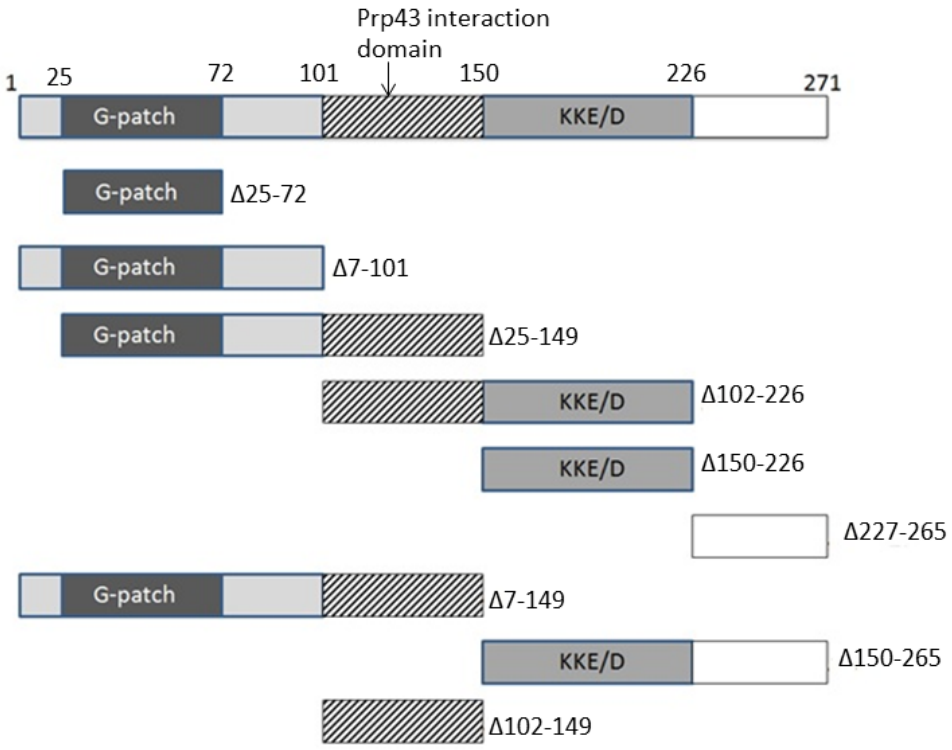

Figure 3.3.1. The domain architecture of Pxr1 (top diagram) and the deletion mutants constructed for the yeast two hybrid assays. The deleted segment is represented as "Δ" followed by the amino acid coordinates.

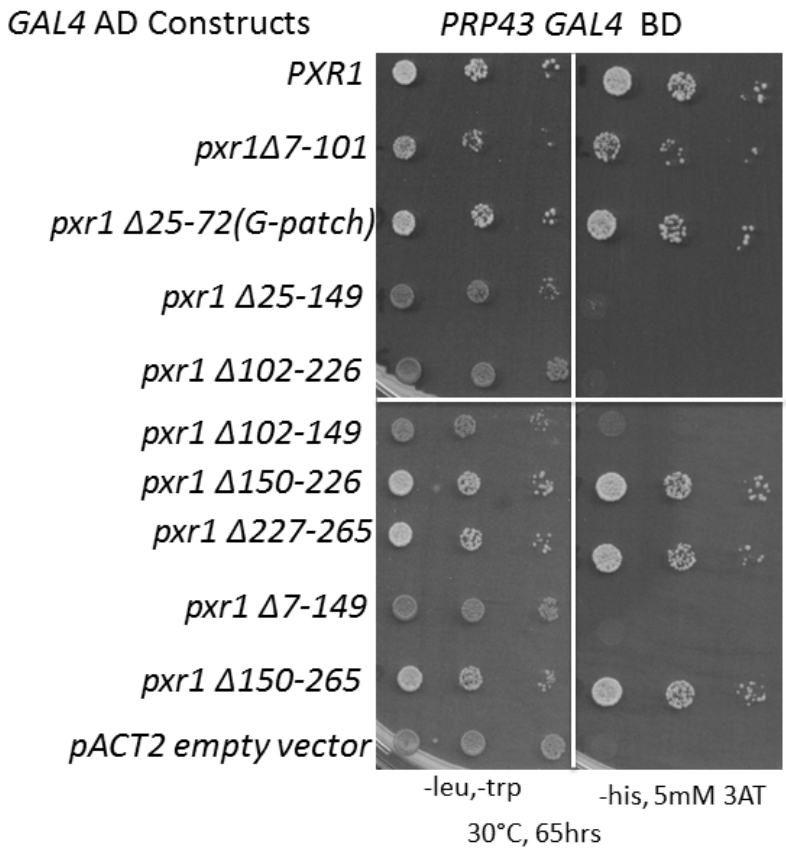

Figure 3.3.2. Removal of a 48 amino acid segment of Pxr1 (102-149 aa) greatly impairs Prp43 binding. Presented are results of the yeast two-hybrid assay with full length Prp43 and deletion mutants of Pxr1 illustrated in figure 3.1. The selective medium for trans-activation includes 5 mM 3-aminotriazole to suppress background.

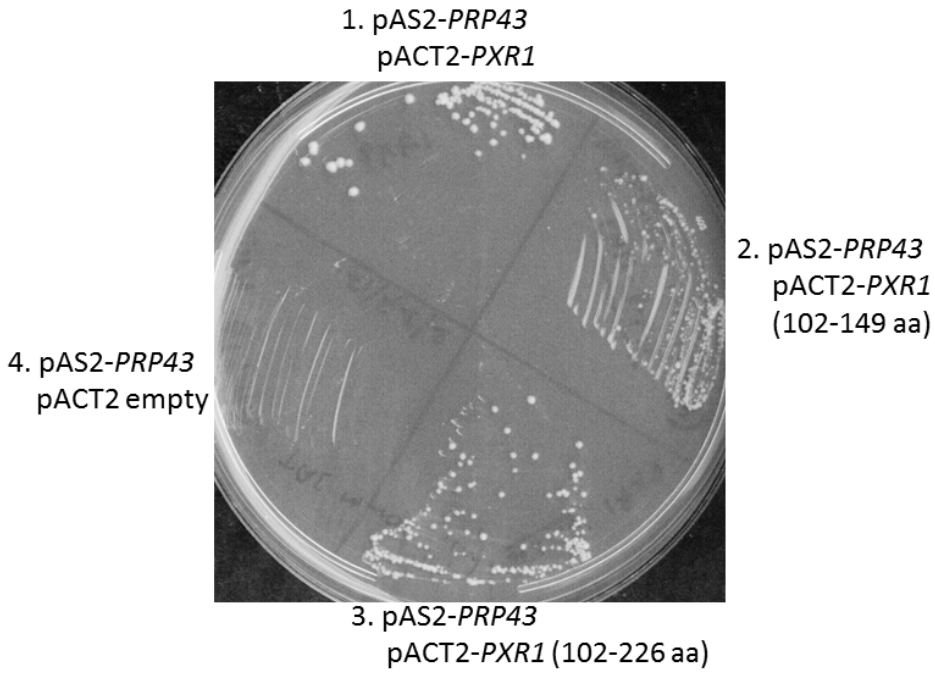

Figure 3.3.3 Pxr1 segment 102-149 aa is sufficient for Prp43 interaction. Growth of yeast on a Y2H trans-activation plate lacking histidine and supplemented with 20 mM 3-amino-triazole. Sector 1 shows interaction of full length Prp43 (Gal4-BD fused) with full length Pxr1 (Gal4-AD fused). Sector 2 and 3 shows interaction of Prp43 with the isolated Pxr1 (102-149 aa) fragment or the Pxr1 (102-226 aa) fragment (Gal4-AD fused) respectively. Sector 4 shows the empty vector negative control. Colonies were photographed after 84 hours at 30°C.

3.3.1 The Pxr1 G-patch, but not its primary Prp43 binding domain is critical for Pxr1 function *in vivo*.

The G-patch of Pxr1 appears fully dispensable for Prp43 interaction in the Y2H assay (Figure 3.2) whereas the region between amino acids 102-149 is critical for this association (Figure 3.3.2 and Figure 3.3.3). We next wanted to learn if the 102-149 Prp43-binding domain is important for the natural Pxr1 biological function. To test this, I first cloned the wildtype *PXR1* gene with its natural promoter into the yeast centromeric plasmid, Ycplac111 (Gietz and Sugino 1988). This plasmid and a series of *pxr1* deletion derivatives were constructed by inverse PCR were then scored for complementation of a viable but strongly growth impaired *pxr1::KAN* mutant (Wach, Brachat et al. 1994, Winzeler, Shoemaker et al. 1999).

As expected, the cloned wildtype *PXR1* gene fully complements the slow growth phenotype of the *pxr1::KAN* mutant, whereas an empty vector control shows no growth improvement (Figure 3.3.4). A G-patch deleted *pxr1* derivative also fails to complement the null allele, consistent with the poor growth observed with a previously described G-patch point mutant (Guglielmi and Werner 2002). Surprisingly, however, the 102-149 domain deletion derivative complements the *pxr1::KAN* null allele almost as well as the wildtype *PXR1* gene, showing that this site of Prp43 interaction is not critical for Pxr1 function. This assay was repeated at 25°C and 37°C and produced equivalent results (data not shown). I note that while the 102-149 and 150-226 deletions support efficient growth, the combined 102-226 aa deletion derivative does fail to complement the *pxr1::KAN* null allele. This suggests that either the independent domains contribute

redundant function or that the larger deletion may render the protein unstable or improperly folded.

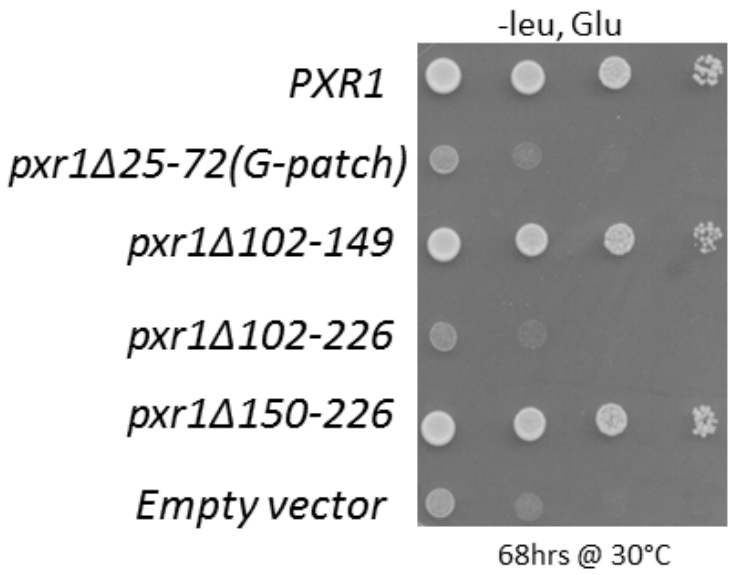

Figure 3.3.4. The newly defined Prp43-binding site of Pxr1 is dispensable for Pxr1 function. This image shows the complementation of *pxr1::KAN* by the wildtype *PXR1* gene or the indicated deletion mutant alleles. The cell cultures were spotted as 10 fold serial dilutions on the indicated selective media and incubated for 68 hours at 30°C.

3.3.2 Pre-rRNA processing efficiency correlates with the cellular growth of mutant Pxr1 harboring strains.

The *pxr1::KAN* mutant complemented with various *PXR1* deletion derivatives (*pxr1ΔG-patch, pxr1Δ102-149, pxr1Δ102-226,* and *pxr1Δ150-226)* showed differences relative colony growth (Figure 3.4). I wanted to test if the growth phenotype correlates with the rRNA processing activity in these mutants. The rRNA maturation pathway was investigated by Northern blot hybridized with ^{32}P-labelled oligonucleotides designed to detect specific rRNA intermediates (Kos and Tollervey 2005), shown by red circles in Figure 3.3.5 . I used the intronless *ADE3* gene probe as a normalization control.

The *pxr1* null mutant containing an empty vector showed an increased accumulation of 35S precursor over what is seen in the transformant bearing the wildtype *PXR1* gene (Figure 3.3.6). The removal of the Pxr1 G-patch resulted in a similar increased accumulation of the 35S precursor, demonstrating that the Pxr1 G-patch is critical for the efficient rRNA processing activity of this protein, consistent with the previous results obtained with a *pxr1* G-patch point mutant (Guglielmi and Werner 2002) and the growth assay presented above.

The null *pxr1* strain complemented with the 102-149 aa segment deleted Pxr1 appears efficient in rRNA processing with no obvious increased 35S pre-rRNA or rRNA intermediate accumulation, suggesting that this segment which appears critical for Prp43 interaction is in fact dispensable for Pxr1 function. Like the 102-149 aa segment, the removal of the Pxr1 150-226 aa segment also does not cause any obvious increased 35S precursor accumulation , suggesting either these segments have redundant function or is dispensable for Pxr1 function. In contrast, removal of the Pxr1 102-226 aa segment leads

to accumulation of 35S precursor, which is consistent with its inability to complement the growth defect of the null *pxr1* mutant (Figure 3.3.4).

In addition to the 35S rRNA precursor, several previously described rRNAs intermediates (27SA2, 23S, 20S) and the mature RNAs (25S, 18S and 5.8S) were assayed for changes in rRNA processing efficiency in the *pxr1* mutant backgrounds. The most obvious change is the decrease 5.8S rRNA abundance in the *pxr1* null, *pxr1ΔG-patch*, *pxr1Δ102-226* mutants when compared with the wildtype Pxr1 strain. Thus, overall the growth phenotype observed in the *pxr1* mutants correlates with the pre-rRNA processing efficiency, indicating that the impaired growth of the *pxr1* strains is predominantly due to impaired ribosomal RNA processing.

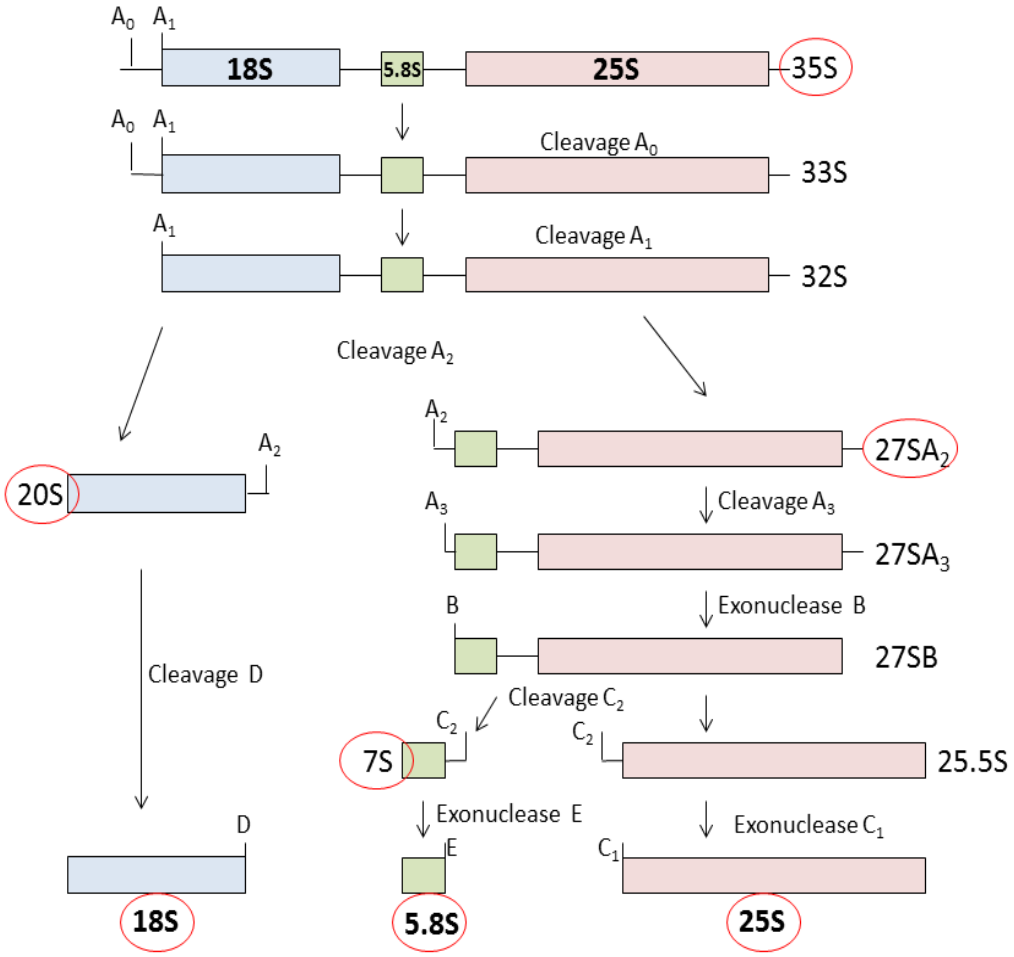

Figure 3.3.5. A schematic representation of the yeast rRNA processing pathway. The 35S precursor ribosomal RNA is processed by endonucleolytic and exonucleolytic cleavages into mature 25S, 18S and 5.8S rRNA. The intermediates steps detected by my Northern blot analysis are indicated by red circles. Image adapted from (Pertschy, Schneider et al. 2009).

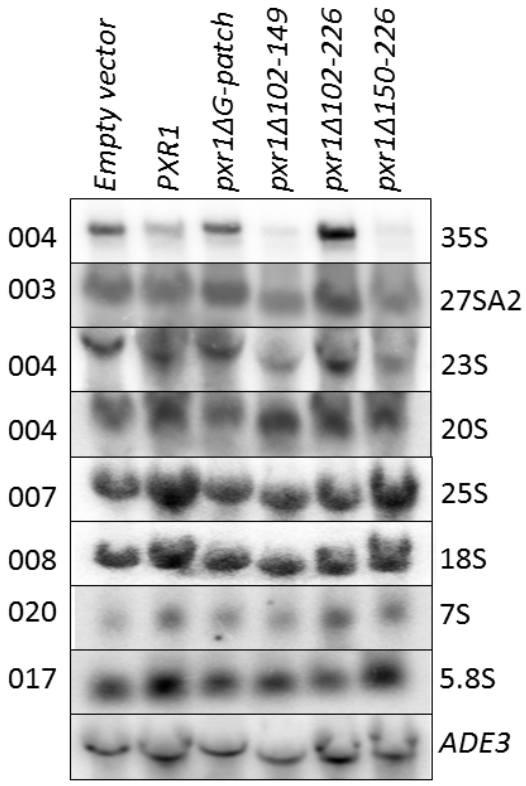

Figure 3.3.6 Analysis of rRNA processing intermediates of various Pxr1 deletion mutants. Northern blot analysis of the steady state pre-rRNA and rRNA levels in the mutant *pxr1* strains. 20 µg of each RNA is resolved on a 1.2% formaldehyde/agarose gel and transferred to a nitrocellulose membrane. The blot was sequentially probed with ^{32}P labeled oligonucleotides, exposed to phosphor screen, with the phosphoimager bands quantified by ImageQuant (see Materials and methods). *ADE3* was used as normalization control. Probes are indicated on the right and the pre-rRNA and the rRNA species are indicated on the left. For quantification data see Appendix Table A1.

3.4 The Sqs1 G-patch reconstitutes Spp382 interaction with Prp43 while the Pxr1 G-patch cannot.

The loss of the Spp382 G-patch blocks Y2H interaction with Prp43 (Figure 3.2). The isolated G-patches of Spp382, Sqs1 and Pxr1 all interact with full length Prp43 by this assay, although these differ in apparent relative affinity (i.e., Spp382>Sqs1>Pxr1; Figure 1). We hypothesized that the G-patch might function as an interchangeable Prp43 binding surface. To test this, I did a G-patch domain swap experiment illustrated in Figure 3.4.1 and scored the chimeric Spp382 constructs for Prp43 interaction.

The G-patch was first deleted from the Gal4-AD fused *SPP382* construct by inverse PCR and then replaced with the cognate *SPP382* G-patch module or with the *PXR1* or *SQS1* G-patch modules; the reconstituted constructs were subsequently scored for interaction with Prp43 in an Y2H assay (Figure 3.4.2). As expected, the Spp382 G-patch efficiently reconstitutes Spp382-Prp43 interaction; the Sqs1 G-patch chimera also restores growth but not quite as strongly (see, *spp382ΔG+SQS1G*). In contrast, the Pxr1 G-patch cannot substitute Spp382 G-patch for Prp43 interaction (*spp382ΔG+PXR1G*). Note, however, that the somewhat reduced colony size on the -leucine, -tryptophan medium (i.e., *spp382ΔG+PXR1G*, left panel) suggests a modest dominant negative impact of this construct in the absence of reporter gene expression.

To provide a more quantitative readout of this data, I scored the *GAL7-lacZ* reporter gene in this strain for β-galactosidase activity and present this data in Figure 3.4.2. The add-back of the cognate Spp382 G-patch shows roughly half of the activity of the unmanipulated Spp382 with Prp43 suggesting that the SacII restriction site incorporated at the site of add-back insertion may be slightly inhibitory. The add-back of

the Sqs1 G-patch shows roughly 33% activity of the cognate Spp382 add-back, whereas the G-patch deleted Spp382 and the Pxr1 G-patch add-back construct does not show activity above background. So, at least in the Y2H assay, substitution of at least one rRNA processing G-patch domain (i.e., that of Sqs1) can restore Prp43 interaction to any otherwise defective Spp382 lacking its G-patch domain.

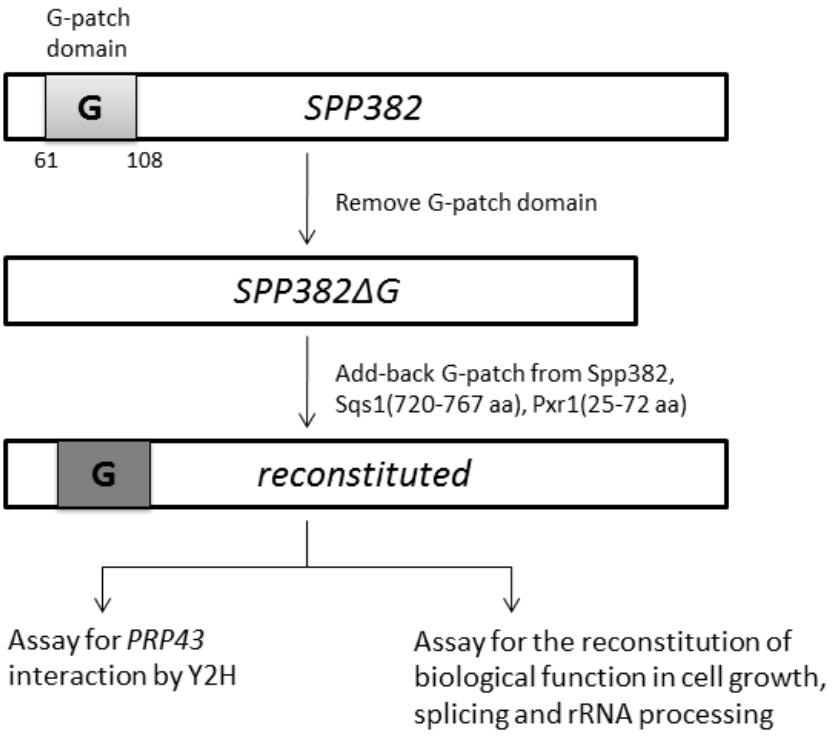

Figure 3.4.1. Outline of the domain swap experiment.

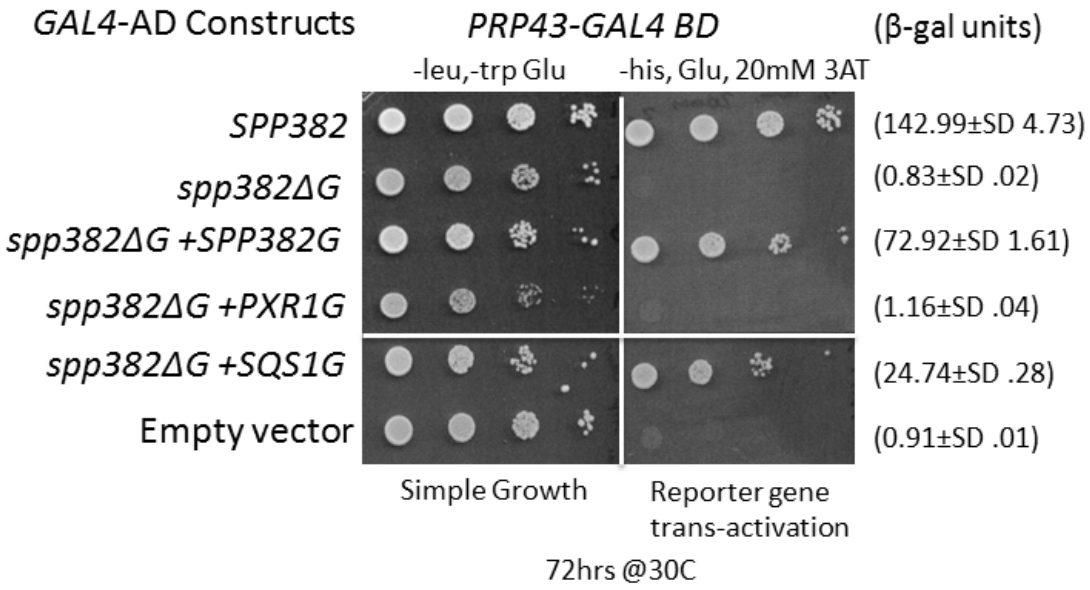

Figure 3.4.2. Spp382-Prp43 Y2H interaction can be supported by the Sqs1 G-patch but not by the Pxr1 G-patch. The Y2H interaction of full length Prp43 and the Spp382 –G patch chimeras. The cultures were incubated at 30°C for 72 hours. *Gal7-LacZ* reporter gene transactivation was also scored with the β-galactosidase units shown in parenthesis.

3.4.1 The Sqs1 G-patch reconstitutes Spp382 activity *in vivo* while the Pxr1 G-patch cannot.

The Sqs1 G-patch supports Spp382 -Prp43 interaction in the Y2H assay but the Pxr1 G-patch cannot (Figure 3.4.2) suggesting related but non-identical activities for the different G-patch motif members. Since it is unclear if G-patch function is limited to Prp43-binding, I next addressed whether the Sqs1 G-patch or the Pxr1 G-patch can reconstitute Spp382 biological function *in vivo*. To do this, I used a "plasmid shuffle" approach (Boeke, Trueheart et al. 1987) to score for function of the chimeric constructs in the previously described yeast strain, N19, which contains a chromosomal *spp382::KAN* null allele complemented by a galactose induced functional copy of *SPP382* on a *URA3*-marked plasmid (*p416-GAL1*::*spp382-4*) (Pandit, Lynn et al. 2006). Inverse PCR was used to remove the *SPP382* G-patch coding sequence from an otherwise wildtype *SPP382* allele expressed on the yeast centromeric plasmid, Ycplac111(Gietz and Sugino 1988). This plasmid expresses *SPP382* with its natural promoter (Pandit, Lynn et al. 2006). Into this *spp382*ΔG construct I inserted either the cognate *SPP382* G-patch, the *PXR1* G-patch or the *SQS1* G-patch sequence and scored each for *spp382::KAN* complementation.

In Figure 3.4.3, the left panel shows growth where the functional *p416-GAL1*::*spp382-4* allele is expressed on galactose media; the right panel shows growth dependent exclusively on the chimeric *SPP382* alleles on FOA media (Boeke, Trueheart et al. 1987). As expected, expression of an unmanipulated wild type *SPP382* gene supports growth in the *spp382::KAN* null allele background while expression of an empty vector *(EMPTY)* or the removal of Spp382 G-patch from *SPP382* (*spp382ΔG*) is lethal,

reinforcing the importance of the G-patch in function (Pandit, Lynn et al. 2006, Tanaka, Aronova et al. 2007). Insertion of the cognate Spp382 G-patch into the *spp382ΔG* background (*spp382ΔG-patch+SPP382G*) restores wildtype growth while insertion of the heterologous Sqs1 G-patch (*spp382ΔG-patch+SQS1G*) supports growth, although clearly less efficiently than observed with the cognate G-patch. On the other hand, Pxr1 G-patch substitution (*spp382ΔG-patch+PXR1 G*) does not support growth showing that this peptide is unable to reconstitute Spp382 function (see Discussion for potential explanations).

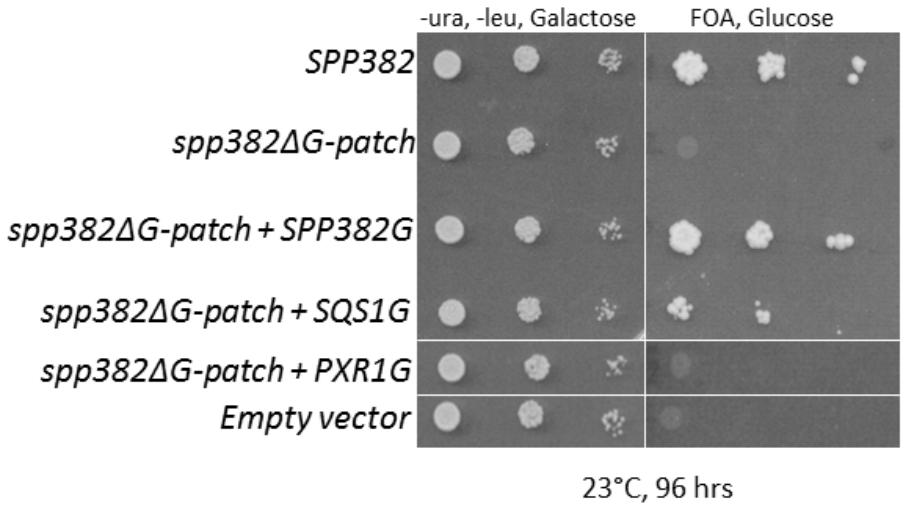

Figure 3.4.3. The Sqs1 G-patch but not the Pxr1 G-patch reconstitutes Spp382 function *in vivo*. Growth of serial dilutions of the N19 transformant yeast cultures (*spp382::KAN*, *p416-GAL1*::*spp382-4)* on the indicated medium. The left panel shows growth on galactose medium where the functional *p416-GAL1*::*spp382* allele is expressed. The right panel shows colony formation on FOA medium where growth is dependent on the chimeric *SPP382* construct. This growth assay was repeated at 30°C yielding identical results, data not shown.

The yeast strain bearing the chimeric Spp382 as the sole source of this essential splicing factor supports cell growth (Figure 3.4.3). The differential growth rate of the chimeric Spp382-Sqs1G-patch transformant compared to the unmanipulated Spp382 and to the cognate Spp382 G-patch add-back strains might be attributed to altered splicing efficiencies. To address this possibility, a probe for the intron containing mRNA for the small subunit ribosomal protein, *RPS17A,* was used in Northern analysis to monitor splicing efficiency which is quantified as a ratio of the processed message (M) to the unprocessed pre-mRNA (P) (Rymond, Pikielny et al. 1990) (Figure 3.4.4A). The intronless *ADE3* mRNA serves as a loading control. Splicing is efficient in yeast bearing the unmanipulated *SPP382* allele, and as expected, in the transformant containing the cognate Spp382 G-patch inserted into the *spp382ΔG*-construct with each showing a mRNA/pre-mRNA in the range of 20-30 previously reported for this gene (Rymond, Pikielny et al. 1990). In contrast, the yeast strain bearing the chimeric Spp382-Sqs1derivative construct shows decreasing two to three fold decrease in the message/precursor ratio, thus consistent with the reduced growth of this viable strain. This suggests the growth defect results from impaired pre-mRNA splicing.

The yeast strains harboring the *spp382ΔG*-patch and *spp383ΔG-patch+PXR1 G-patch* chimeras are not viable in the absence of the functional *SPP382* copy (Figure 3.4.3). The splicing efficiency of the *spp382ΔG-patch* and the *spp382ΔG-patch+PXR1 G-patch* chimera bearing strains were assessed by the transcriptional repression of a functional Spp382 from *GAL1*::*spp382-4* (Pandit, Lynn et al. 2006). As expected, the *spp383ΔG-patch* and the *spp383ΔG-patch+PXR1 G-patch* mutants are both greatly splicing impaired, and leads to accumulation of the unprocessed precursor (Figure 3.4.4B). This

shows that the Spp382 G-patch domain is critical for its role in splicing and the Pxr1 G-patch fails to compensate for its loss.

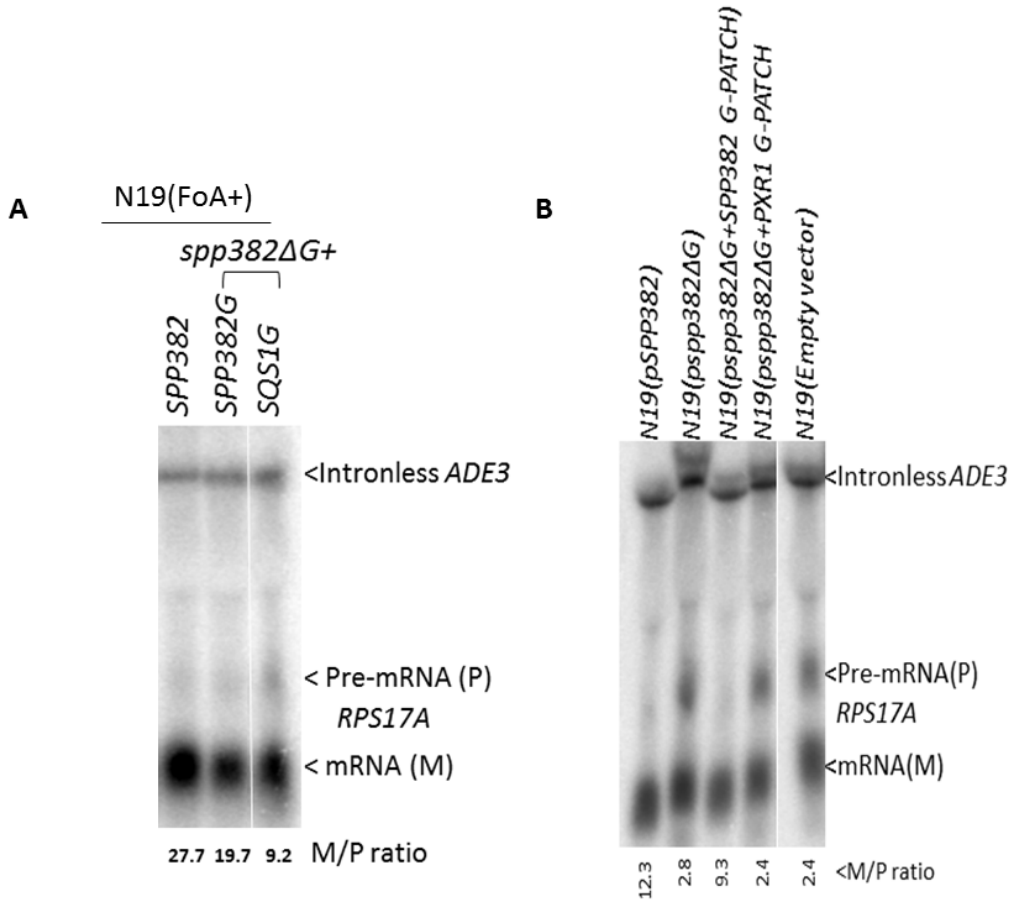

Figure 3.4.4A. The Spp382-Sqs1 G-patch chimera shows reduced splicing efficiency. Northern analysis of the RNA isolated from yeast strains with the indicated Spp382 constructs as the sole source of the protein. **B. Pxr1 G-patch fails to compensate for the Spp382 G-patch's role in splicing.** Northern analysis of the RNA isolated from yeast with the indicated Spp382 contructs after transcriptional repression of the functional *GAL1*::*spp382-4* gene by growth for 18 hours in glucose media. The positions of the loading control *ADE3* mRNA, the intron-bearing *RPS17A* precursor (P) and the splicied *RPS17A* message (M) are indicated on the right with arrow-heads.

3.4.2 The Sqs1 G-patch and the Spp382 G-patch can substitute for the Pxr1 G-patch in Pxr1 cellular function.

In the context of the essential spliceosomal factor Spp382, the Sqs1 G-patch but not the Pxr1 G-patch can promote interaction with Prp43 (Figure 3.4.2) and more importantly, support Spp382 function *in vivo* (Figure 3.4.3 and 3.4.4). Based on this, there appears to be overlapping but not identical function among the G-patch domains of pre-mRNA and rRNA processing factors, at least between the G-patches of Sqs1 and Spp382. To pursue this issue further, I investigated the pattern of reconstitution when the Pxr1 G-patch was removed and replaced by a heterologous G-patch. I used inverse PCR to remove the G-patch from the Ycplac111-*PXR1* construct (Figure 3.3.4) and replaced it with either the cognate Pxr1 cassette, the Sqs1 G-patch or the Spp382 G-patch cassette.

As shown above (Figure 3.3.4), a plasmid copy of the wildtype *PXR1* gene fully complements the *pxr1::KAN* null allele *(*Δ*pxr1,PXR1)* while G-patch deletion mutant (Δ*pxr1,pxr1*Δ*G*) fails to complement (Figure 3.4.5). As predicted, replacement of G-patch with the cognate Pxr1 G-patch (Δ*pxr1,pxr11*Δ*G+PXR1G*) fully complements the *pxr1::KAN* null mutation). Interestingly, the Sqs1 G-patch efficiently substitutes for the Pxr1 G-patch domain (Δ*pxr1,pxr1*Δ*G+SQS1G*) while the Spp382 G-patch shows much reduced activity (Δ*pxr1,pxr1*Δ*G+SPP382G*). These data reinforce the modular nature of this domain and once again show differences in the pattern of function – here, the Sqs1 G-patch being more “Pxr1-like” than the Spp382 G-patch.

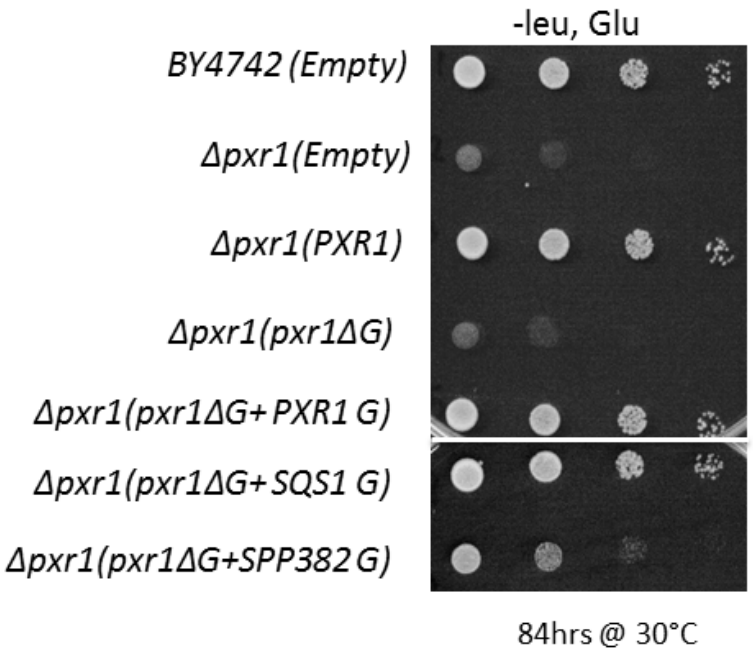

Figure 3.4.5. The Sqs1 G-patch and Spp382 G-patches differentially restore Pxr1 function. Growth of wildtype yeast (*BY4742*) or the yeast *pxr1::KAN* strains (Δ*pxr1*) bearing indicted gene after 84 hours at 30°C.

3.4.3 The Sqs1 or Spp382 G-patch can promote Pxr1 function in rRNA processing.

The *pxr1ΔG-patch* deletion derivative can be reconstituted to support efficient growth by insertion of the Sqs1 G-patch and, to a lesser extent, by the Spp382 G-patch (Figure 3.4.5). To assess the efficacy of ribosomal RNA (rRNA) processing in these chimeric backgrounds, I performed a Northern blot analysis with ^{32}P-labelled oligonucleotides designed to detect specific rRNA intermediates (Kos and Tollervey 2005) and see Figure 3.3.5. I observed an increase in the 35S, 27SA2 and 23S pre-rRNA intermediates in the *pxr1::KAN* mutant background (*pxr1Δ, Empty vector*) over the wild-type *PXR1* strain (Figure 3.4.6, lane 8 and 9), consistent with earlier findings by (Guglielmi and Werner 2002). The 25S rRNA and, somewhat less obviously, the 18S levels are also reduced in this mutant and it shows increases in the 35S, 27SA2 and 23S pre-rRNA intermediates. A decrease in the 5.8S rRNA is observed in the *pxr1::KAN* strain and in the *pxr1ΔG-patch* strain. A slight decrease of the 20S rRNA levels was observed in the *pxr1::KAN* null mutant compared to control although this particular rRNA intermediate showed variability in intensity that did not directly correlate with either the increased 35S, 27SA2 and 23S pre-rRNA intermediates or decreases in the 25S rRNA. I could not convincingly detect the 32S precursor that was reported to be modestly depleted in the *pxr1*::KAN null background by (Guglielmi and Werner 2002). The Pxr1 G-patch deletion mutant (lane 10) behaved roughly the same as the untransformed *pxr1::KAN* mutant and resulted in an increased accumulation of 35S pre-rRNA, 27SA2 rRNA, 23S pre-rRNA intermediates and reductions in the 7S, 5.8S, 18S and 25S RNA species.

Similar to what was seen in the growth assay, the add-back of cognate Pxr1 G-patch (lane 11) restored rRNA processing to the level seen in the unmanipulated wild-type strain. Likewise, based on the relative band intensities of the rRNA intermediates and mature RNA products, the Sqs1 G-patch showed much better improvement in rRNA processing (lane 13) than what was observed with the Spp382-G-patch (lane 12) although rRNA processing was less still efficient than what is seen in the wildtype strain, indicating that neither the Spp382 nor the Sqs1 G-patch does restore complete Pxr1 activity.

In addition to the chimera studies conducted to investigate G-patch reconstitution of Pxr1 in rRNA processing, we demonstrated earlier that the Sqs1 G-patch can reconstitute Spp382 activity sufficient enough to restore growth, while the Pxr1 G-patch substitution cannot (Figure 3.4.3). The chimeric constructs supported pre-mRNA splicing, albeit less well than the wildtype Spp382 (Figure 3.4.4). Since these chimeras fuse domains from established pre-mRNA splicing and rRNA splicing factors, it is possible that cross-pathway interference occurs and may contribute to the reduced growth rates of the chimeric strains. To address this, I scored the viable chimeric constructs for rRNA processing defects (Figure 3.4.6). I found no obvious differences in rRNA processing between the Spp382 –Sqs1G-patch chimera (lane 6) and wildtype (lane 3 and 7). The 5.8 S rRNA band intensity is lower in the Spp382-Sqs1 G-patch construct as well as in the severely growth impaired culture where no Spp382 is expressed (empty vector, lane 1). At this point it is unclear whether this difference reflects a specific defect in rRNA processing attributable to this Spp382-Sqs1 G-patch construct or other factors that influence sample recovery.

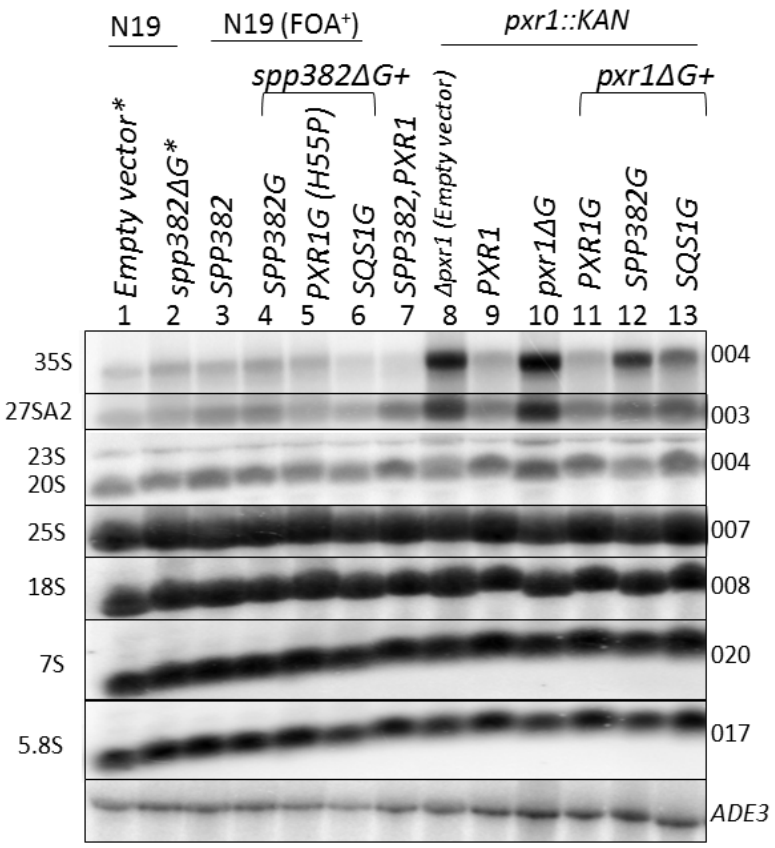

Figure 3.4.6. Sqs1 or Spp382 G-patch promotes Pxr1 rRNA processing activity *in vivo*. Northern analysis of the steady state pre-rRNA and rRNA in the chimeric Pxr1 (lanes 11-13) and chimeric Spp382 strains (lanes 4-6) are analyzed. RNA was harvested from the indicated strains and resolved on a 1.2% formaldehyde/agarose gel and transferred to a nitrocellulose membrane. The blot was sequentially probed with ^{32}P labeled oligonucleotides, exposed to phosphor screen, and the phosphoimager bands were quantified by ImageQuant (see Materials and methods). *ADE3* was used as normalization control. Probes are indicated on the right and the pre-rRNA and the rRNA species are indicated on the left. "*" in lanes 1 and 2 indicates that the *spp382ΔG* construct and control were expressed after repression of a *GAL1*-driven functional copy. For quantification data see Appendix Table A2.

3.5 Identification of G-patch identity features required for Spp382 reconstitution.

The data presented above establish that the Spp382, Pxr1 and Sqs1 G-patches do not act equivalently in reconstituting Spp382 function in either the Y2H Prp43 interaction assay (Figure 3.4.2) or in complementing the lethal *spp382::KAN* mutation (Figure3.4.3 and 3.4.4) or in equivalent assays for Pxr1 function (Fig 3.4.4 and 3.4.5). Presumably the amino acid composition of each G-patch is optimized for its natural biological role and this role differs between the G-patch proteins involved in splicing and rRNA processing. As a first step to identifing amino acid residues that confer functional specificity, I did a multiple sequence alignment of the Spp382, Pxr1 and Sqs1 G-patch domains across species for each family. The putative G-patch protein homologs were identified by a protein-protein BLASTP (BLOSUM 62 matrix) search in NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) with the full length *S. cerevisiae* G-patch protein against the non-redundant (nr) protein databases of *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Arabidopsis thaliana, Drosophila melanogaster, Caenorhabditis elegans, Danio rerio, Xenopus laevis, Gallus gallus, Mus musculus, Bos Taurus* and *Homo sapiens*). The matches with the lowest E-value were retrieved from National Center for Biotechnology Information (NCBI) Protein Database (http://www.ncbi.nlm.nih.gov/protein/) using accession number provided in Table 2.2 (see Materials and methods). The G-patch coordinates for each protein were taken from NCBI protein database.

Multiple Sequence Alignment was performed with the G-patches of Spp382 (Sp), Pxr1 (Px) and Sqs1 (Sq) across 11 species using MUSCLE (Edgar 2004, Edgar 2004) (http://www.ebi.ac.uk/Tools/msa/muscle/) and the results are presented in Figure 3. 5.

The general G-patch consensus sequence (Aravind and Koonin 1999) is shown on top, followed by the consensus sequences of the Spp382, Pxr1 and Sqs1 G-patches across the eleven species studied that were generated by MUSCLE alignment. As anticipated I observe in the G-patch family consensus sequences the previously reported conserved glycine residues (positions 7, 15, 19, 21,23, 30 and 48; coordinates refer to the number grid on the top line), hydrophobic amino acids (positions 10, 11 and 22) and aromatic amino acid (position 16) common to the generic G-patch motif (Aravind and Koonin 1999).

In addition to the G-patch consensus core positions, I found examples of residues phylogenetically conserved within a protein family grouping (same in Spp382 G-patch family, or in Sqs1 G-patch family or in Pxr1 G-patch family) but that differ between families (some examples shown with asterisk below the alignments in Figure 3. 5). Such residues are candidates for amino acids that confer G-patch specificity to the Spp382, Pxr1 and Sqs1-like proteins.

I hypothesize that the G-patch residues differing between Spp382 and Pxr1 account for the poor performance of the Pxr1 G-patch in Spp382 reconstitution. To test this, I mutated five such residues in Pxr1 (R27G, P48G, H55P, K57E and D62M, the numbers indicate amino acid coordinate within Pxr1) to the corresponding Spp382 residue to make this G-patch more Spp382-like. Each construct was then scored for the ability to reconstitute Spp382 for Prp43 interaction by the Y2H assay and in complementation of the lethal *spp382::KAN* mutation

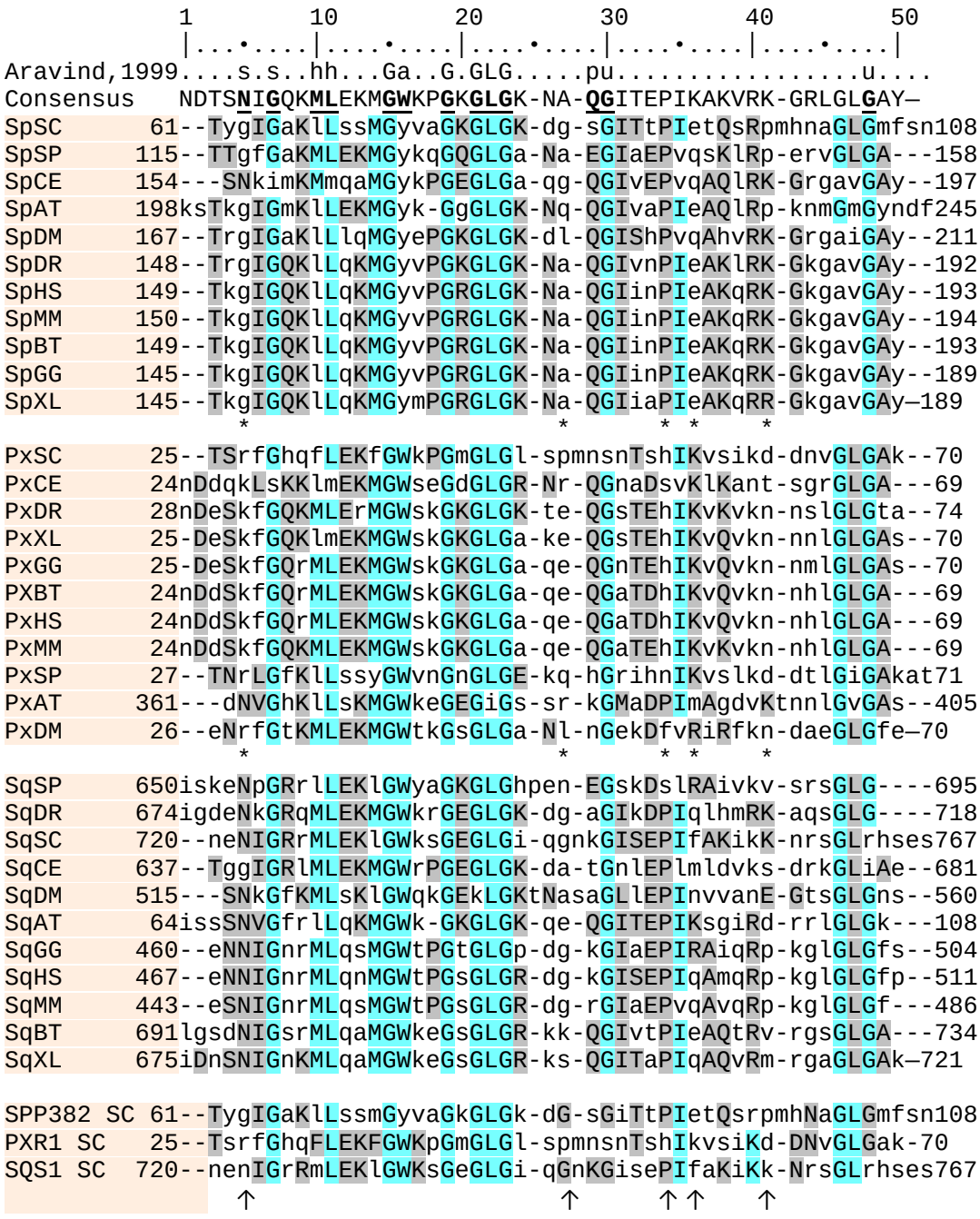
1 10 20 30 40 50
Aravind,1999....s.s..hh...Ga..G.GLG.....pu.....u....
Consensus NDTSNIGQKMLEKMGWKPGKGLGK-NA-QGITEPIKAKVRK-GRLGLGAY—
SpSC 61--TygIGaKlLssMGyvaGKGLGK-dg-sGITtPIetQsRpmhnaGLGmfsn108
SpSP 115--TTgfGaKMLEKMGykqGQGLGa-Na-EGIaEPvqsKlRp-ervGLGA---158
SpCE 154---SNkimKMmqaMGykPGEGLGa-qg-QGIvEPvqAQlRK-GrgavGAy--197
SpAT 198ksTkgIGmKlLEKMGyk-GgGLGK-Nq-QGIvaPIeAQlRp-knmGmGyndf245
SpDM 167--TrgIGaKlLlqMGyePGKGLGK-dl-QGIShPvqAhvRK-GrgaiGAy--211
SpDR 148--TrgIGQKlLqKMGyvPGKGLGK-Na-QGIvnPIeAKlRK-GkgavGAy--192
SpHS 149--TkgIGQKlLqKMGyvPGRGLGK-Na-QGIinPIeAKqRK-GkgavGAy--193
SpMM 150--TkgIGQKlLqKMGyvPGRGLGK-Na-QGIinPIeAKqRK-GkgavGAy--194
SpBT 149--TkgIGQKlLqKMGyvPGRGLGK-Na-QGIinPIeAKqRK-GkgavGAy--193
SpGG 145--TkgIGQKlLqKMGyvPGRGLGK-Na-QGIinPIeAKqRK-GkgavGAy--189
SpXL 145--TkgIGQKlLqKMGymPGRGLGK-Na-QGIiaPIeAKqRR-GkgavGAy—189
PxSC 25--TSrfGhqfLEKfGWkPGmGLGl-spmnsnTshIKvsikd-dnvGLGAk--70
PxCE 24nDdqkLsKKlmEKMGWseGdGLGR-Nr-QGnaDsvKlKant-sgrGLGA---69
PxDR 28nDeSkfGQKMLErMGWskGKGLGK-te-QGsTEhIKvKvkn-nslGLGta--74
PxXL 25-DeSkfGQKlmEKMGWskGKGLGa-ke-QGsTEhIKvQvkn-nnlGLGAs--70
PxGG 25-DeSkfGQrMLEKMGWskGKGLGa-qe-QGnTEhIKvQvkn-nmlGLGAs--70
PXBT 24nDdSkfGQrMLEKMGWskGKGLGa-qe-QGaTDhIKvQvkn-nhlGLGA---69
PxHS 24nDdSkfGQrMLEKMGWskGKGLGa-qe-QGaTDhIKvQvkn-nhlGLGA---69
PxMM 24nDdSkfGQKMLEKMGWskGKGLGa-qe-QGaTEhIKvKvkn-nhlGLGA---69
PxSP 27--TNrLGfKlLssyGWvnGnGLGE-kq-hGrihnIKvslkd-dtlGiGAkat71
PxAT 361---dNVGhKlLsKMGWkeGEGiGs-sr-kGMaDPImAgdvKtnnlGvGAs--405
PxDM 26--eNrfGtKMLEKMGWtkGsGLGa-Nl-nGekDfvRiRfkn-daeGLGfe—70
SqSP 650iskeNpGRrlLEKlGWyaGKGLGhpen-EGskDslRAivkv-srsGLG----695
SqDR 674igdeNkGRqMLEKMGWkrGEGLGK-dg-aGIkDPIqlhmRK-aqsGLG----718
SqSC 720--neNIGRrMLEKlGWksGEGLGi-qgnkGISEPIfAKikK-nrsGLrhses767
SqCE 637--TggIGRlMLEKMGWrPGEGLGK-da-tGnlEPlmldvks-drkGLiAe--681
SqDM 515---SNkGfKMLsKlGWqkGEkLGKtNasaGLlEPInvvanE-GtsGLGns--560
SqAT 64issSNVGfrlLqKMGWk-GKGLGK-qe-QGITEPIKsgiRd-rrlGLGk---108
SqGG 460--eNNIGnrMLqsMGWtPGtGLGp-dg-kGIaEPIRAiqRp-kglGLGfs--504
SqHS 467--eNNIGnrMLqnMGWtPGsGLGR-dg-kGISEPIqAmqRp-kglGLGfp--511
SqMM 443--eSNIGnrMLqsMGWtPGsGLGR-dg-rGIaEPvqAvqRp-kglGLGf---486
SqBT 691lgsdNIGsrMLqaMGWkeGsGLGR-kk-QGIvtPIeAQtRv-rgsGLGA---734
SqXL 675iDnSNIGnKMLqaMGWkeGsGLGR-ks-QGITaPIqAQvRm-rgaGLGAk—721
SPP382 SC 61--TygIGaKlLssmGyvaGkGLGk-dG-sGiTtPIetQsrpmhNaGLGmfsn108
PXR1 SC 25--TsrfGhqFLEKFGWKpGmGLGl-spmnsnTshIkvsiKd-DNvGLGak-70
SQS1 SC 720--nenIGrRmLEKlGWKsGeGLGi-qGnKGisePIfaKiKk-NrsGLrhses767

Figure 3.5. Multiple Sequence Alignment of G-patch domains Across Species.

Alignment was done using MUSCLE program (Edgar 2004, Sonnhammer and Hollich 2005). A given column has one color which indicates the average BLOSUM62 score: light blue ≥ 3, dark blue ≥ 1, light gray ≥ 0.2, no color otherwise. The general G-patch consensus from (Aravind and Koonin 1999) is shown on top (h-hydrophobic (LIYFWVMA); l-aliphatic (LIVAM); a-aromatic (FYW); s-small (GASNSTCP); u-tiny (GAS) and p-polar (STNREQHD)). Numbers at the beginning and end of each sequence denotes the approximate G-patch co-ordinates for each protein obtained from NCBI protein database. The core G-patch consensus sequence (Aravind and Koonin 1999) (glycine, hydrophobic and aromatic amino acids) are bold and underlined in the consensus of Spp382 (Sp), Pxr1(Px) and Sqs1 (Sq) G-patches . Examples of family specific identical residues are indicated by asterisk "*". Residues in *S.cerevisiae* Pxr1 G-patch mutated to corresponding Spp382 G-patch residues are shown in arrow heads. Abbreviations: *Saccharomyces cerevisiae*- SC , *Schizosaccharomyces pombe*- SP, *Arabidopsis thaliana*- AT, *Drosophila melanogaster*- DM, *Caenorhabditis elegans*- CE, *Danio rerio*- DR, *Xenopus laevis*- XL, *Gallus gallus*- GG, *Mus musculus*- MM, *Bos Taurus*- BT and *Homo sapiens*- HS.

3.6 Making the Pxr1 G-patch more Spp382-like enhances Spp382-Pxr1 chimera interaction with Prp43 in the Y2H assay.

We hypothesize that amino acid differences within the G-patch motif of Pxr1 and Spp382 account for the poor performance of the Spp382-Pxr1 G-patch chimera in the Y2H assay and in complementing the *spp382::KAN* null allele (Figure 3.4.2 and 3.4.3). To test this, I scored the Spp382-like Pxt1 G-patch mutants (R27G, P48G, H55P, K57E and D62M) ability to reconstitute Spp382-Prp43 interaction (Figure 3.6.1). The right panel shows simple growth of the doubly transformed yeast and the left panel shows growth on medium requiring transactivation of *GAL-HIS3* reporter gene. As reported before, Spp382 interacts with Prp43 and deletion of the Spp382 G-patch blocks this interaction (Tsai, Fu et al. 2005, Tanaka, Aronova et al. 2007) and see Figure 3.3.2. The Pxr1 G-patch cannot substitute Spp382 G-patch (*spp382*Δ*G+PXR1 G*) in this assay. Also, Spp382-Pxr1 chimeras with G-patch changes R27G and H55P fail to show Y2H interaction above background. In contrast, Pxr1 G-patch mutations P48G, K57E and D62M each promote Spp382- Prp43 interaction when assayed by this method at 23°C. This appears to be a comparatively weak interaction based on the slow growth observed at 23°C and the fact that the Y2H interaction is lost at 30°C (data not shown). Note, however, that the somewhat reduced colony size on the -leucine, -tryptophan medium (i.e., *spp382ΔG+PXR1G, P48G, K57E, D62M* left panel) suggests a modest dominant negative impact of this construct in the absence of reporter gene expression.

To quantify the Y2H data, I also scored these constructs for transactivation of a second host reporter gene, *GAL7-lacZ*. Table 3.1 shows the β-galactosidase activities for each strain. The wildtype Spp382 and Prp43 interaction yields a significantly increased β-

galactosidase activity as compared to the pACT2 empty vector control. The removal of Spp382 G-patch domain abolishes this interaction whereas addition of cognate Spp382 restores β-galactosidase activity to nearly full activity. The slightly reduced activity in the Spp382 G-patch add-back chimera compared with the unmanipulated gene might be attributed to interference resulting from the two SacII sites introduced in the add-back construct. The add-back of the Pxr1 G-patch or the Pxr1 R27G construct does not significantly enhance β-galactosidase activity above background whereas Pxr1 G-patch derivatives P48G, H55P, K57 and D62M each show modest but statistically significant 2 to 2.4-fold increase in activity consistent with the growth results shown with the transactivation of the *GAL1-HIS3* reporter.

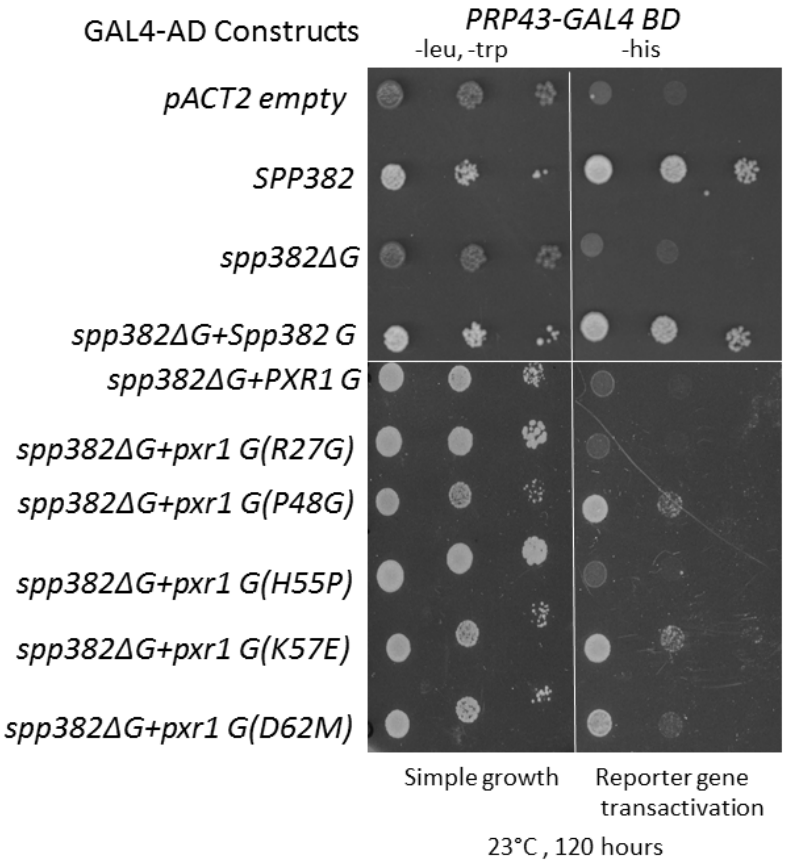

Figure 3.6.1. Chimeric Spp382 with select Spp382-like mutated Pxr1 G-patch cassettes modestly enhance Prp43 interaction. Yeast 2 hybrid assay of full length Prp43 (Gal4-BD fused) and chimeric Spp382 with wild-type or mutant Pxr1 G-patch (Gal4-AD fused) is shown as a colony growth assay, where each in each column is spotted 10 fold serial dilutions of the indicate cultures incubated at 23°C for 120 hours. The positions of the point mutations within the Pxr1 protein coding sequence are indicated in parenthesis.

pJ69-4a (pAS2-*PRP43*) with pACT2-	β-Gal activity (Miller units) n=3	P value (Unpaired t-test)
Empty	1.17±SD.14	
SPP382	320.91±SD11.53	<0.0001*
spp382ΔG	1.18±SD.19	0.945
spp382ΔG+SPP382 G	199.14±SD25.73	0.0002*
spp382ΔG+PXR1 G	1.25±SD.10	0.4657
spp382ΔG+pxr1 G (R27G)	1.07±SD.14	0.4311
spp382ΔG+ pxr1 G(P48G)	2.44±SD.13	0.0003*
spp382ΔG+ pxr1 G(H55P)	1.9±SD.17	0.0045*
spp382ΔG+ pxr1 G(K57E)	2.19±SD.12	0.0006*
spp382ΔG+ pxr1 G(D62M)	2.02±SD.22	0.0049*

Table 3.1. β-galactosidase activity of yeast strain pJ69-4a harboring the indicated pAS2-Prp43 and pACT2-Spp383 chimeric Pxr1 G-patch constructs. β-galactosidase activities measured as described by (Miller 1972, Mockli and Auerbach 2004). For statistical significance unpaired t-test was conducted using the website: http://www.graphpad.com/quickcalcs/ttest1.cfm. "*" indicates statistical significance, n represents the number of replicates.

3.6.1 Select Pxr1 G-patch mutations reconstitute Spp382-Pxr1 chimera biological activity *in vivo* but these do not necessarily correlate with increased Y2H interaction.

Unlike the Spp382-Pxr1 (WT) construct, the chimeric Spp382 constructs containing the mutant Pxr1 G-patch cassettes (P48G, H55P, K57 and D62M) show low but detectable Prp43 interaction by the Y2H assay. We were interested to know if the increased Prp43 interaction correlated with a change in the ability of these chimeric proteins to reconstitute Spp382 cellular function. To address this, these same chimeric constructs were prepared in the context of an otherwise wildtype *SPP382* gene and scored for complementation of the lethal *spp382::KAN* mutation (Pandit, Lynn et al. 2006, Tanaka, Aronova et al. 2007). Since one or more of these constructs may be non-functional, the chimeric genes were introduced as plasmid copies into the N19 yeast strain described above which contains a second plasmid that expresses the galactose responsive and biologically active *GAL1::spp382-4* allele (Pandit, Lynn et al. 2006). After selecting the double transformants based on the plasmid-linked nutrituional markers, I used the plasmid shuffle technique (Sikorski and Boeke 1991) to score for the ability of the chimeric constructs to support cell viabilty in the absence of *GAL1::spp382-4*.

Shown in Figure 3.6.2 is an image of the plasmid shuffle assay on FOA media of the N19 strains bearing the Spp382 chimeria with the original (wildtype) Pxr1 G-patch, mutant Pxr1 G-patches or an empty vector control. Four out of five single amino acid substitutions (i.e., R27G, H55P, K57 and D62M) support cellular viability, only the

chimeric Spp382 construct bearing the wildtype Pxr1 G-patch and the P48G Pxr1 G-patch are lethal.

The viable strains that express the chimeric Spp382- Pxr1 proteins as the sole source of this essential splicing factor were next plated on minimal media to compare relative growth rate by the culture dilution assay (Figure 3.6.3). *spp382ΔG+pxr1 G-patch (H55P)* supports cell growth most efficiently, forming colonies roughtly half the size of the unmanipulated wildtype strain. The *spp382ΔG+pxr1 G-patch (R27G), spp382ΔG+pxr1 G-patch (K57E)* and *spp382ΔG+pxr1 G-patch (D62M)* derivatives are viable but, as seen above on the FOA plate (Fig. 3.6.2), grow poorly.

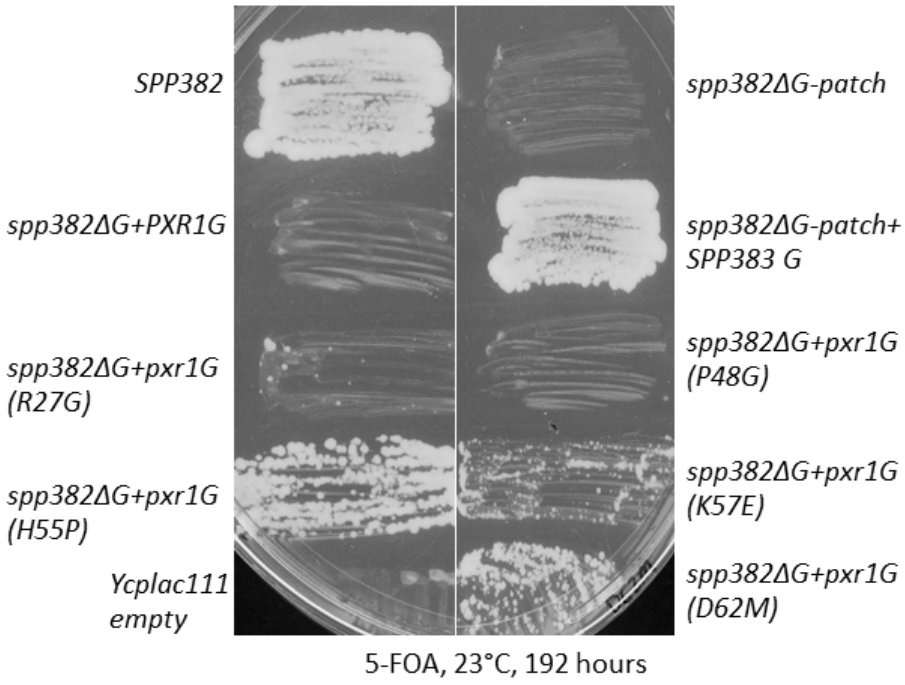

Figure 3.6.2. Certain *spp382-pxr1* G-patch mutant constructs complement the lethal *spp382::KAN* allele. A plasmid shuffle assay of *spp382::KAN* yeast that contain a *URA3*-linked functional *GAL1::spp382-4* plasmid and a *LEU2*-linked second plasmid bearing with and without the indicated *SPP382-PXR1* chimeric contracts streaked on 5-fluoroorotic acid (FOA) medium to select against the *GAL1::spp382-4* plasmid. The results show that Pxr1 G-patch derivatives R27G, H55P, K57E and D62M are all capable of supporting cell viability as Spp382-Pxr1 chimeras.

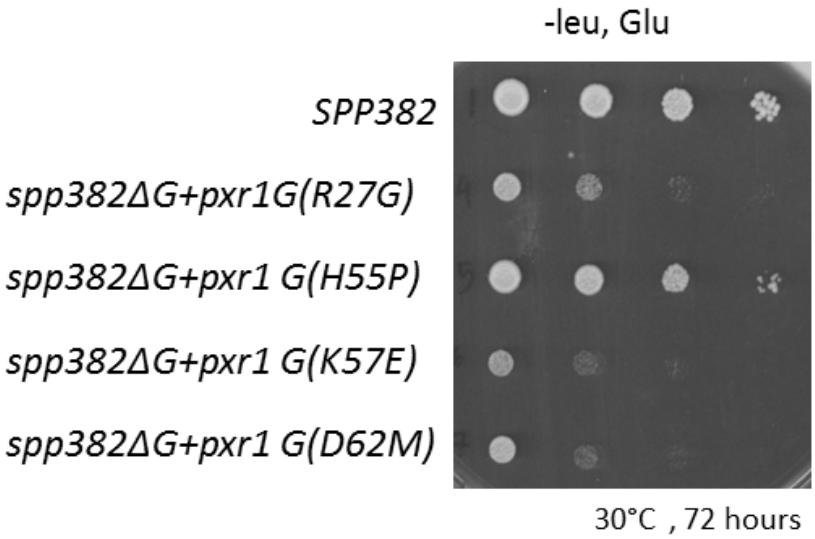

Figure 3.6.3. The Spp382-Pxr1 (H55P) chimera efficiency supports cellular growth most efficiently. Growth of yeast dependent solely on the chimeric *SPP382* construct (shown on left) spotted on the indicated medium as a 10 fold dilution series.

3.6.2 Pre-mRNA splicing efficiency correlates with the cellular growth of chimeric Spp382 harboring strains.

The viable yeast strains bearing chimeric Spp382-Pxr1G-patch mutants as the sole source of this essential splicing factor support cell growth to varying degrees with the *spp382ΔG+pxr1 G-patch (H55P)* transformants growing well and the *spp382ΔG+pxr1 G-patch (R27G), spp382ΔG+pxr1 G-patch (K57E)* and *spp382ΔG+pxr1 G-patch (D62M)* transformants forming slow growing colonies (Figure 3.6.3). The differential growth rate of these chimeric Spp382 bearing yeast strains might be attributed to varied splicing efficiencies. To address this possibility, a probe for the intron containing mRNA for the small subunit ribosomal protein, *RPS17A*, was used in a Northern analysis to monitor splicing efficiency as quantified by the ratio of the processed message (M) to the unprocessed pre-mRNA (P) (Rymond, Pikielny et al. 1990) (Figure 3.6.4). The intronless *ADE3* mRNA serves as a normalizing loading control.

Splicing is efficient in yeast bearing the unmanipulated *SPP382* allele, and as expected, in the transformant containing the cognate Spp382 G-patch inserted into the *spp382ΔG*-construct. Overall, the yeast strains bearing chimeric the Spp382-Pxr1 derivative constructs show a two to five-fold decrease in message/precursor ratios, consistent with impaired cellular pre-mRNA splicing (Rymond, Pikielny et al. 1990). However, the splicing efficiency of the *spp382-pxr1 (H55P)* chimera is reproducibly better than the other chimeric Spp382 constructs, consistent with the better growth of this yeast strain.

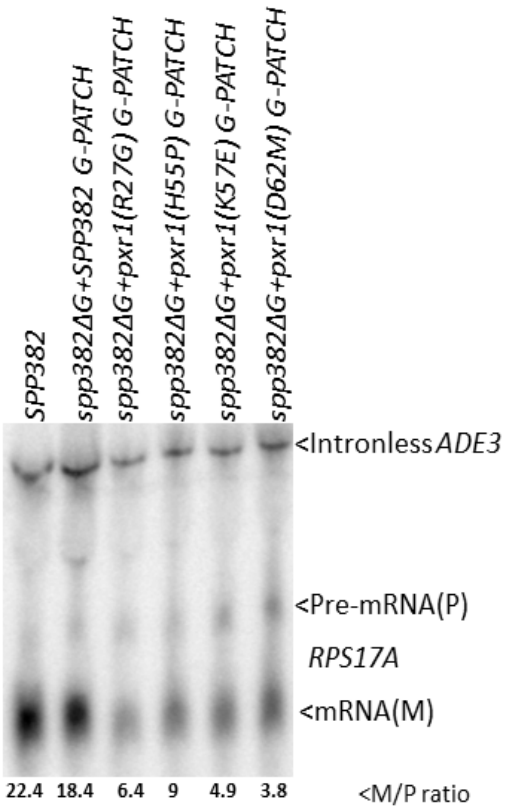

Figure 3.6.4. The *spp382-pxr1 (H55P)* construct shows the greatest splicing efficiency among the Spp382 chimeras. Northern analysis of the RNA isolated from the chimeric Spp382 harboring yeast strains. The positions of the loading control, *ADE3*, the *RPS17A* precursor (P) and message (M) are indicated on the right with arrow-heads. The mRNA/pre-mRNA ratios are indicated at the bottom of each lane. The band intensities were quantified with ImageQuant(GE) software.

3.6.3 Pre-rRNA processing efficiency is somewhat impaired with the expression of certain Spp392-Pxr1 chimeras.

The yeast strains harboring the chimeric Spp382 with mutant Pxr1 G-patches as the sole source of cellular Spp382 clearly exhibit reduced growth on minimal media (Figure 3.6.3) and impaired splicing efficiency compared to a wildtype strain (Figure 3.6.4). However, in addition to the pre-mRNA splicing defects, it is conceivable that adding a domain from an established ribosomal RNA processing factor to Spp382 inhibits rRNA processing. To address this, I did Northern blot analysis of the chimeric Spp382 bearing yeast strains and investigate the rRNA maturation pathway (Figure 3.6.5). I observe an accumulation of the 35S pre-rRNA in the strains expressing the weakest chimeras (supporting growth and splicing complementation), namely *spp382ΔG-patch+pxr1 G-patch (K57E)* and *spp382ΔG-patch+pxr1 G-patch (D62M)* compared to the strain expressing wildtype Spp382. In addition, the 5.8S levels in the *spp382ΔG-patch+pxr1 G-patch (K57E)* appears reduced and there is a generally lower level of the 27SA2 pre-rRNA intermediate in the strains expressing the chimeras. Of these strains, yeast that express the K57E and D62M chimeras grew most slowly and this correlates well with the unique enrichment of the 35S rRNA precursor in these strains.

Thus, the differential ability of the chimeric Spp382 to support growth correlates with reduced efficiency of both pre-mRNA splicing and ribosomal RNA processing. Note, the reduced colony size with the expression of some chimeric *spp382-pxr1 G-patch* constructs (Figure 3.6.1, left panel) in the Y2H strain PJ69-4a. This modest dominant negative impact might be due to interference of the chimeric protein with the fully functional Spp382. As the mature 25S and 18S rRNAs appear largely unchanged, while

splicing is strongly compromised, I suspect that the severely impaired growth observed with K57E and D62M chimeras in the *spp382::KAN* background results largely to impaired Spp382 function in pre-mRNA splicing.

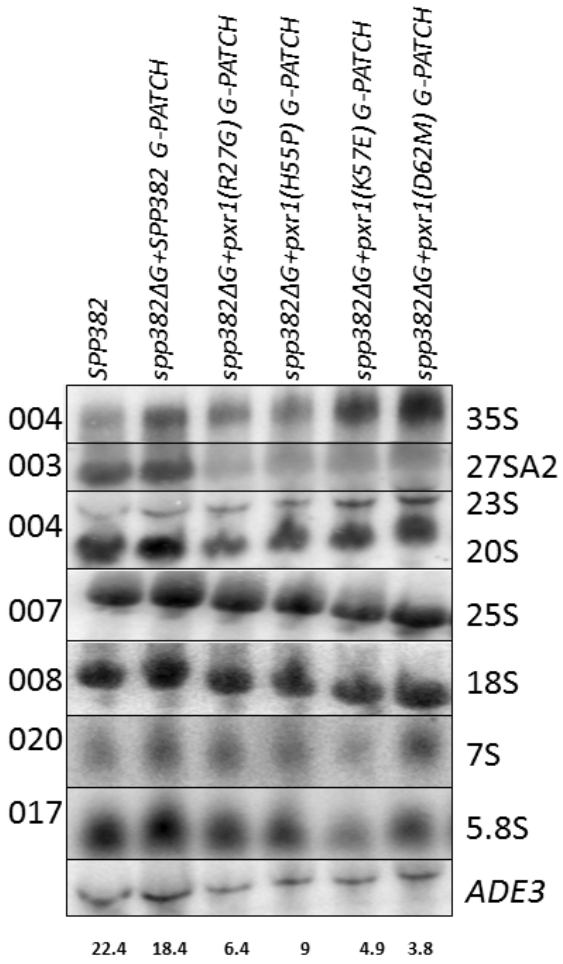

Figure 3.6.5. Varied ribosomal RNA processing efficiency in the *spp382-pxr1* chimeras. rRNA processing efficiency studied by Northern blot of total RNAs isolated from yeast that express the chimeric Spp382-Pxr1 constructs. ^{32}P- labeled oligonucleotides (listed on the left) were used to identify specific rRNA precursor, intermediates, and mature products (listed on the left). *ADE3* hybridization was used as a loading control. For quantification data see Appendix Table A3.

3.7 Identification of the G-patch binding site within Prp43.

Prp43 has two essential biological functions, to promote rRNA processing and to promote pre-mRNA splicing. It physically interacts with three G-patch domain bearing proteins, the rRNA processing factors Pxr1 and Sqs1 and the splicing factor Spp382 as well as other proteins active in both RNA processing pathways (Lebaron, Froment et al. 2005, Gavin, Aloy et al. 2006). We were interested in defining the binding sites for the G-patch proteins in Prp43 and especially its binding site for the G-patch motif. Structural studies from the Nielsen group and the Henry group describe six structural domains within the Prp43 protein (He, Andersen et al. 2010, Walbott, Mouffok et al. 2010). Guided by these studies, I first created a series of Prp43 domain deletion derivatives (Figure 3.7.1) in the yeast 2-hybrid construct pAS2-*PRP43* (Gal4 DNA binding domain fused Prp43) and scored each for interaction with full length G-patch proteins (Figure 3.7.2, 3.7.3 and 3.7.4) or isolated G-patches cassettes (Figure 3.7.8, 3.7.9 and 3.7.10).

3.7.1 Removal of the Prp43 helicase domains RecA1 and RecA2 and the helicase associated winged-helix domain (WH) inhibits G-patch protein interaction.

As discussed above (Figure 3.2), full-length Spp382 interacts well with Prp43 by the Y2H assay. Deletion of the N-terminal domain (NTD), oligonucleotide/oligosaccharide binding site containing C-terminal domain (CTD), and Ratchet domain have no detectable impact on interaction with Spp382 showing that none of these protein domains are critical for Spp382-Prp43 binding. Deletion of the helicase-associated winged-helix (WH) domain clearly impairs Spp382 interaction while deletion of either of the conserved helicase domains (RecA1 or RecA2) blocks the Prp43-Spp382 Y2H signal.

Figure 3.7.3 shows yeast-2 hybrid interaction of this same set of Prp43 domain deletion derivatives with full-length Sqs1. As with the Spp382 interaction, deletion of the NTD, CTD or the Ratchet domain of Prp43 does not inhibit interaction with Sqs1. Curiously, removal of the NTD of Prp43 partially ameliorates the cytotoxic effect of Sqs1 expression that is seen with wildtype Prp43 (Figure 3.7.3, *prp43ΔNTD*, left panel compare Empty Vector, PRP43, and *prp43ΔNTD* lanes). Similar to what was seen above for Spp382, removal of the Prp43 RecA1, RecA2 or WH domain abolishes the Sqs1-Prp43 interaction.

The yeast 2 hybrid interaction of full length Pxr1 and the Prp43 domain deletion derivatives are shown in Figure 3.7.4. Here again, similar to what was seen with Spp382, removal of the NTD, CTD or Ratchet domains has little or no impact on the Prp43-Pxr1 response while deletion of the RecA2 domain blocks this response. Unlike the Spp382, deletion of the Prp43 WH domain blocks Pxr1 interaction. Also, removal of the RecA1 helicase domain from Prp43 results in a more modest defect in Prp43 interaction than is seen with either Spp382 or Sqs1.

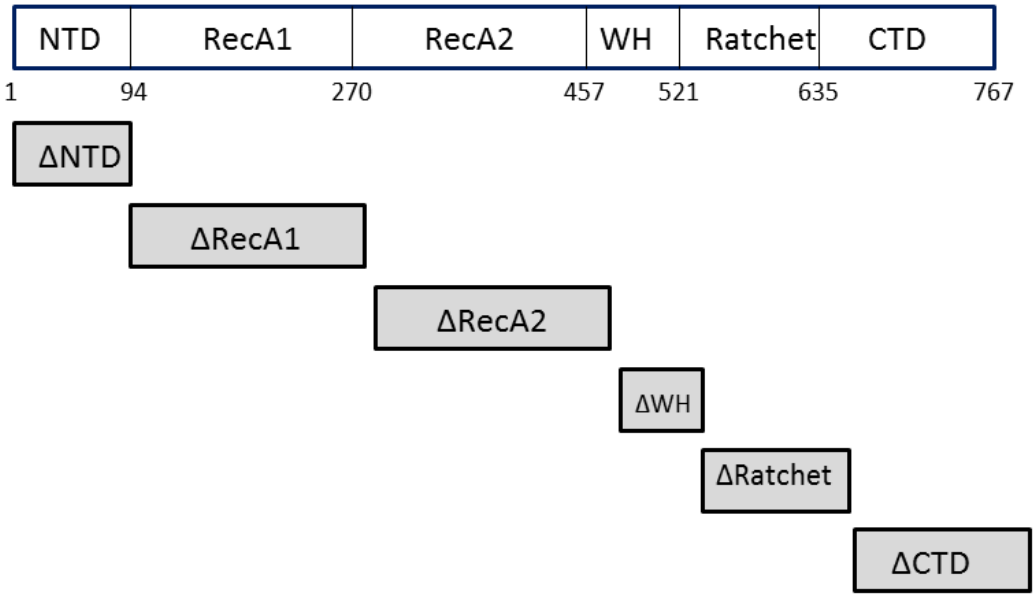

Figure 3.7.1. **The domain architecture of Prp43 and the deletion mutants constructed for the yeast two hybrid assays.** The deleted segment is represented as "Δ". The numbers represent the amino acid positions delimiting the domains.

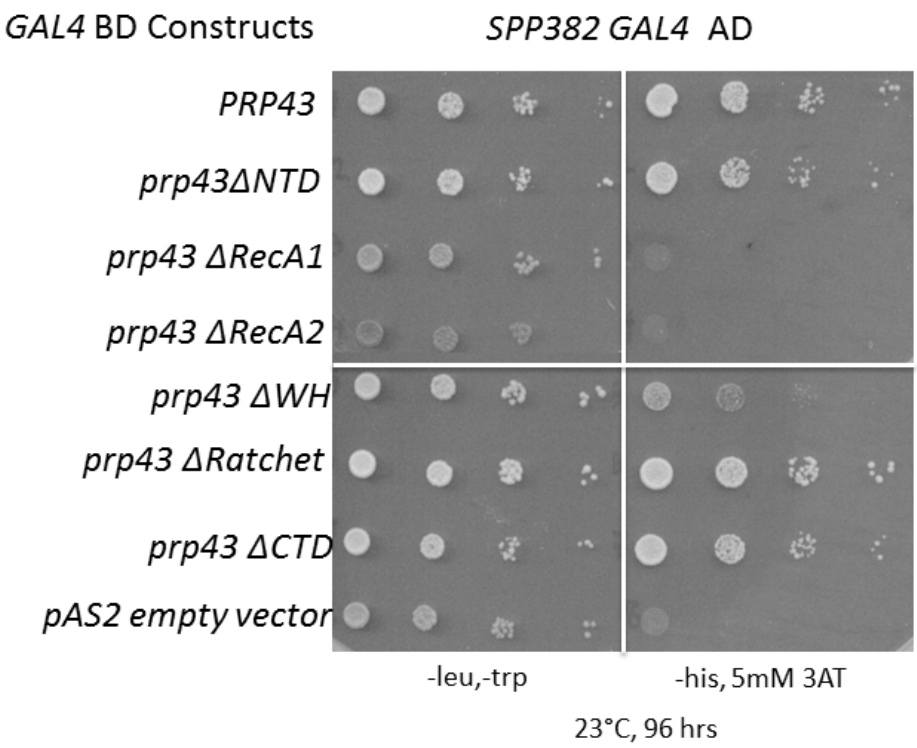

Figure 3.7.2. Deletion of Prp43 helicase domains impairs interaction with Spp382. Yeast-2 hybrid assay of Prp43 deletion domain and full length Spp382 is shown as yeast growth assay. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel shows simple growth to select for yeast-2 hybrid plasmids, right panel shows reporter transactivation.

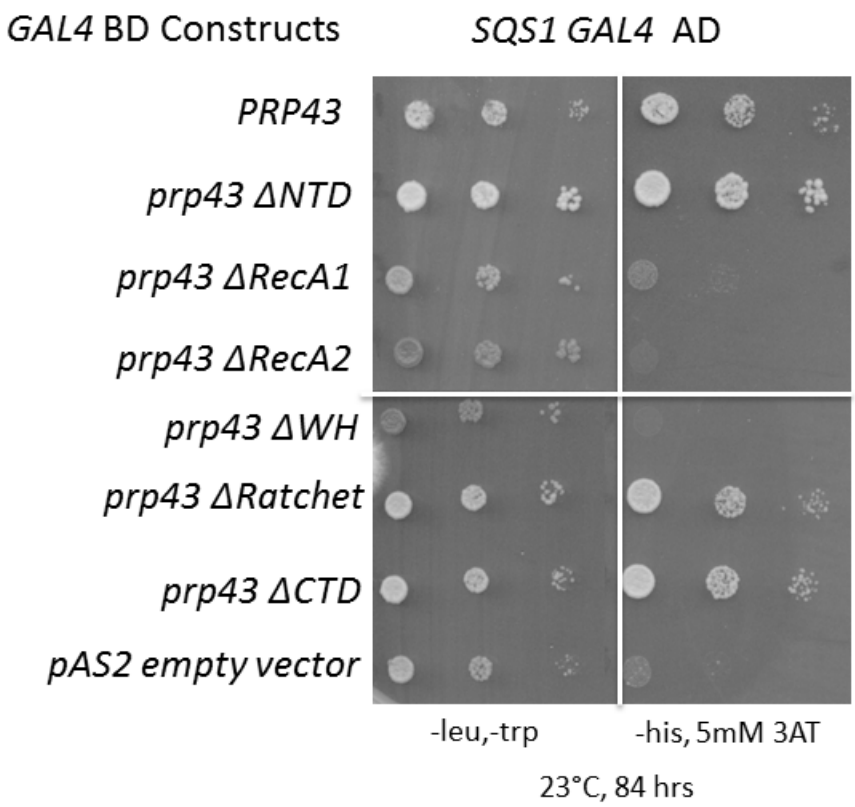

Figure 3.7.3. Deletion of Prp43 helicase domains and helicase associated WH domain impairs interaction with Sqs1. Yeast-2 hybrid assay of Prp43 deletion domain and full length Sqs1 is shown as yeast growth assay. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel shows simple growth to select for yeast-2 hybrid plasmids, right panel shows reporter transactivation.

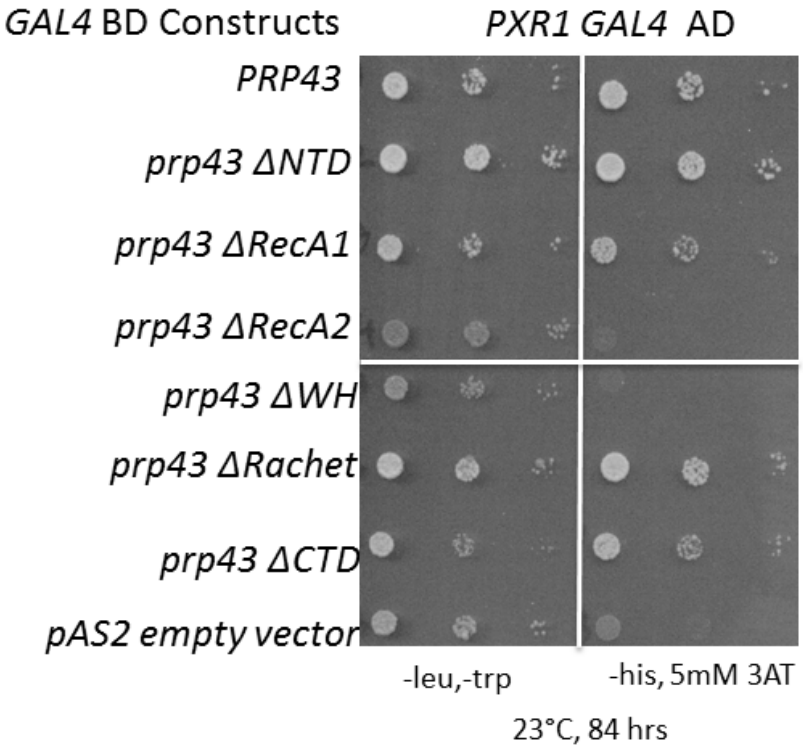

Figure 3.7.4. Deletion of Prp43 RecA2 helicase domain and helicase associated WH domain impairs interaction with Sqs1. Yeast-2 hybrid assay of Prp43 deletion domain and full length Pxr1 is shown as yeast growth assay. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel shows simple growth to select for yeast-2 hybrid plasmids, right panel shows reporter transactivation.

3.7.2 Prp43 Winged-helix (WH) domain is sufficient to interact with G-patch proteins.

The data presented above suggests that the interaction interface of Prp43 with the G-patch proteins is most strongly impaired by deletions of RecA1, RecA2 and, at least for the Pxr1 and Sqs1 proteins, with the removal of the helicase associated winged-helix domain (WH). The loss of G-patch protein interaction might be due to the deletion of a protein binding site or due to the creation of an unstable or improperly folded protein. To begin to address this, I next scored the six isolated Prp43 domains for interaction with each full length G-patch protein.

Detectable interaction is not evidenced by isolated NTD, RecA1, RecA2, Ratchet or CTD (fused to Gal4 DNA binding domain) with full-length Spp382 (Figure 3.7.5). In contrast, the isolated 56 amino acid winged helix domain of Prp43 is capable of interacting with full length Spp382, although with lower activity when compared with the full-length protein.

Pxr1 behaved similarly, with the Prp43 WH domain sufficient for interaction. Surprisingly, the Pxr1-Prp43 WH Y2H interaction is reproducibly stronger than with the full-length Prp43 protein and Pxr1 (Figure 3.7.6). The Prp43 NTD, CTD, RecA1 or Ratchet fail to interact with Pxr1 and Spp382 while, in contrast to Spp382 which is negative in this assay, the isolated RecA2 domain shows weak activity when paired with Pxr1. Sqs1 also shows weak interaction with RecA2 and, similar to Spp382 and Pxr1, a pronounced interaction with the Prp43 WH domain (Figure 3.7.7).

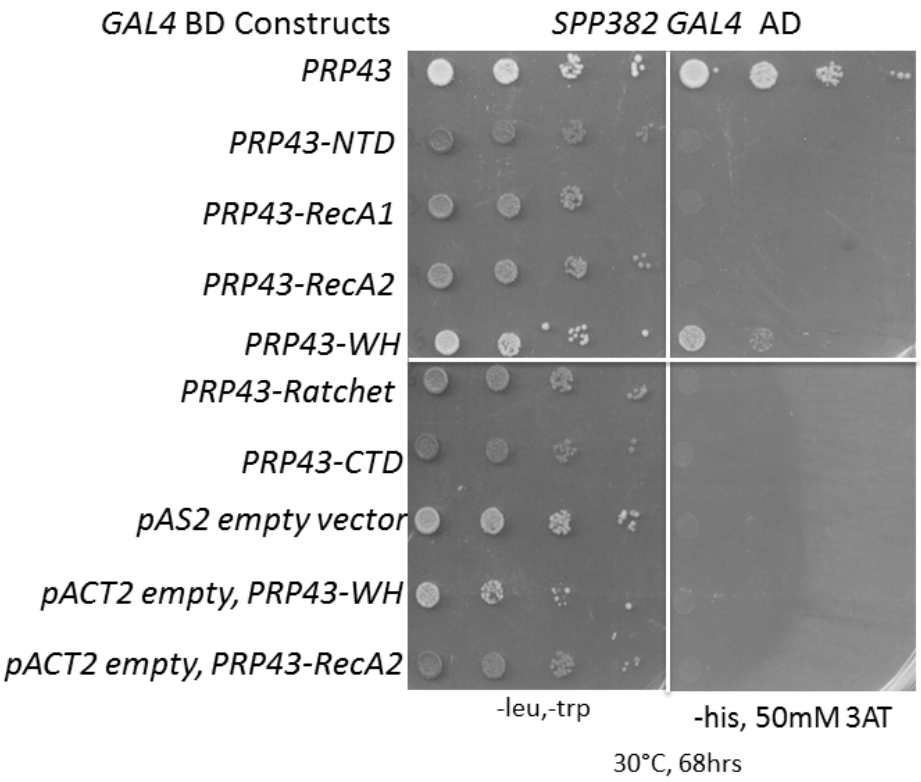

Figure 3.7.5. Prp43 WH domain is sufficient to interact with Spp382. Yeast-2 hybrid assay of Prp43 isolated domains and full length Spp382 is shown as yeast growth assay. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel shows simple growth to select for yeast-2 hybrid plasmids, right panel shows reporter transactivation.

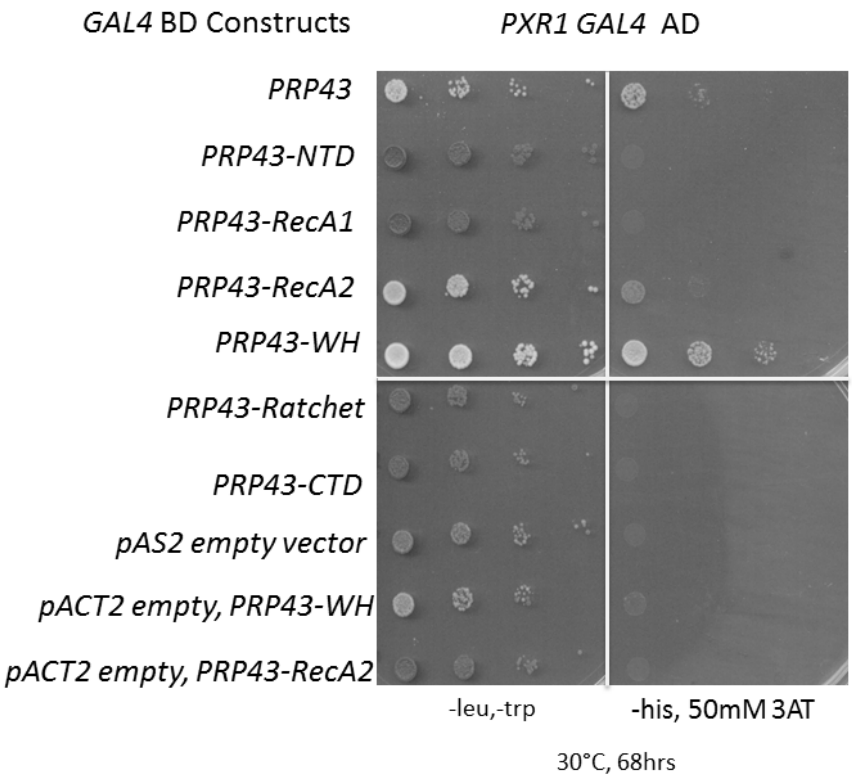

Figure 3.7.6. Prp43 WH domain is sufficient to interact with Pxr1. Yeast-2 hybrid assay of Prp43 isolated domains and full length Pxr1 is shown as yeast growth assay. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel shows simple growth to select for yeast-2 hybrid plasmids, right panel shows reporter transactivation.

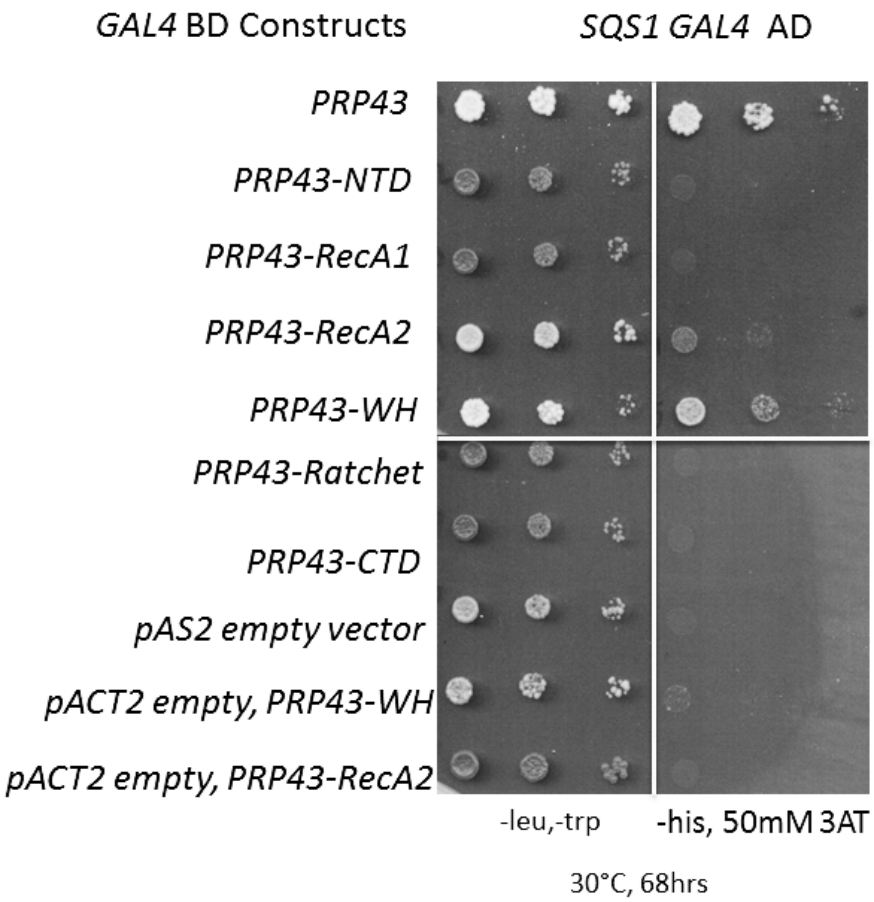

Figure 3.7.7. Prp43 WH domain is sufficient to interact with Sqs1. Yeast-2 hybrid assay of Prp43 isolated domains and full length Spp382 is shown as yeast growth assay. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel shows simple growth to select for yeast-2 hybrid plasmids, right panel shows reporter transactivation.

The yeast 2 hybrid studies with Prp43 deletion derivatives suggest that helicase domains (RecA1 or RecA2) or the helicase-associated WH domain are important for interactions with the G-patch proteins. Interaction studies with isolated Prp43 domains and full length G-patch proteins bolstered these findings and demonstrated that the WH domain is sufficient to interact with all three of the G-patch proteins. I previously showed that each isolated G-patch is sufficient to interact with Prp43, albeit with lower activity than each full length protein (Figure 3.1). Given that the G-patch is the one feature identified in common with Spp382, Pxr1 and Sqs1, I next wanted to determine whether the WH-domain interaction might the common site of association. Figure 3.7.8 shows the Y2H results of interaction of selected Prp43 domain deletion derivatives with the isolated Spp382 G-patch domain. Deletion of RecA1, RecA2or WH domain fails to show detectable activity with Spp382 G-patch domain. The activity of the *prp43ΔRacthet*-Spp382 G-patch pair is significantly reduced compared to the full length Prp43-Spp382 G-patch, although some activity is observed (Appendix 3). Deletion of the Prp43 CTD also fails to show activity above background although this construct interacted well with the full-length Spp382 (Figure 3.7.8).

Extending this study with the Sqs1 and Pxr1 G-patches failed to show detectable interaction above background with Prp43 deletion derivatives in the yeast-2 hybrid analysis under the assay conditions (Figure 3.7.9 and 3.7.10). Strikingly, deletion of the N-terminal Prp43 domain bolstered its interaction with the Spp382 G-patch over that seen with the full length Prp43 protein (Figure 3.7.8). A direct comparison between the pAS2-*prp43ΔNTD* and the three different pACT2- isolated G-patch domains shown in Figure

3.7.11 and Table 7 reveals that the signal with Spp382 G-patch interaction is significantly greater than that seen with either the Pxr1 G-patch or the Sqs1 G-patch.

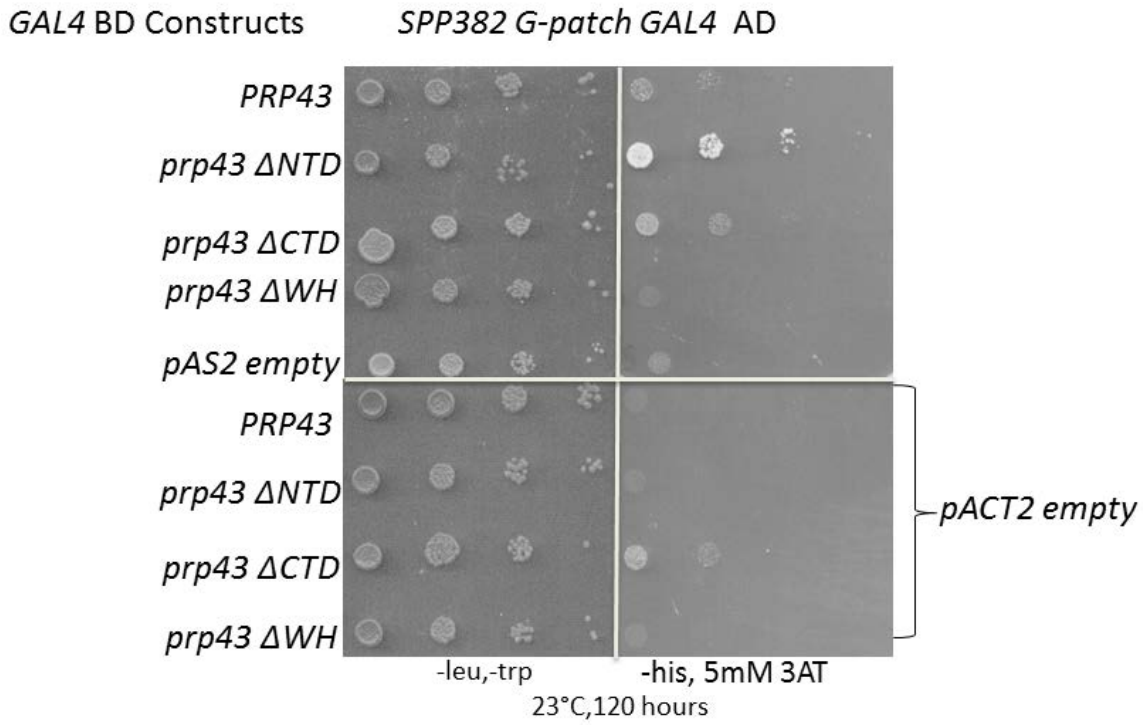

Figure 3.7.8. Removal of Prp43 NTD enhances its interaction with Spp382 G-patch. Yeast-2 hybrid assay of Prp43 deletion derivatives and Spp382 G-patch is shown as yeast growth assay. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel represents simple growth to select for yeast-2 hybrid plasmids, right panel represents reporter transactivation.

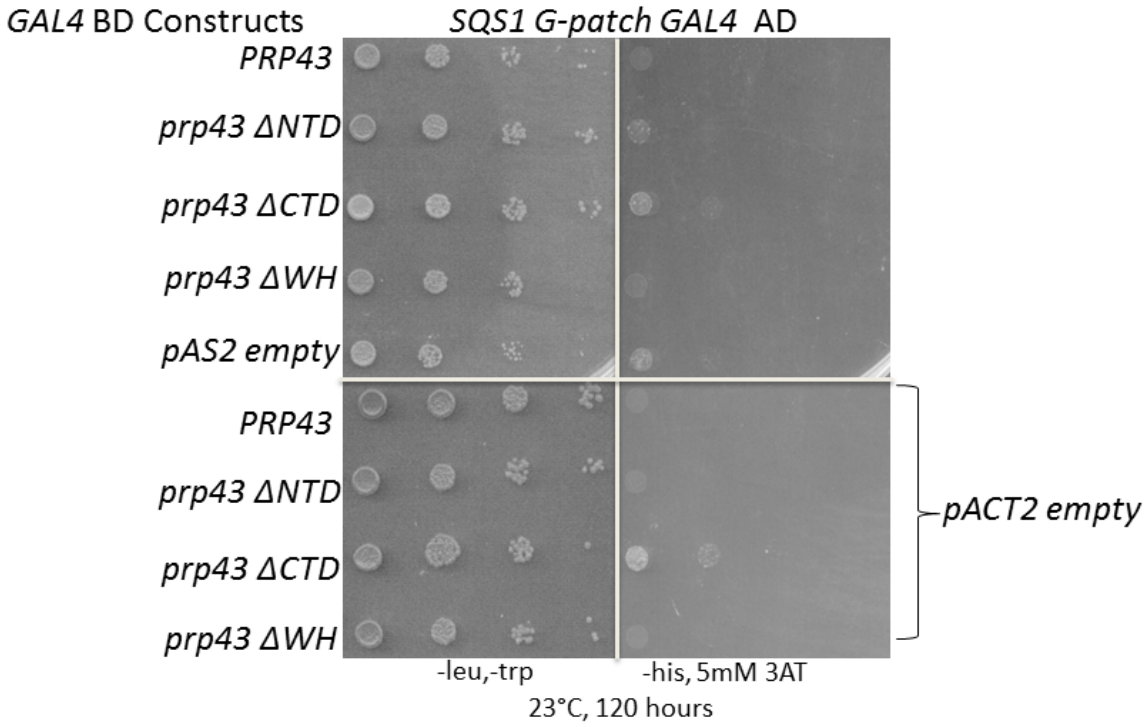

Figure 3.7.9. Removal of Prp43 domains fail to impact interaction with Sqs1 G-patch. Yeast-2 hybrid assay of Prp43 deletion derivatives and Sqs1 G-patch is shown as yeast growth assay. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel shows simple growth to select for yeast-2 hybrid plasmids, right panel shows reporter transactivation.

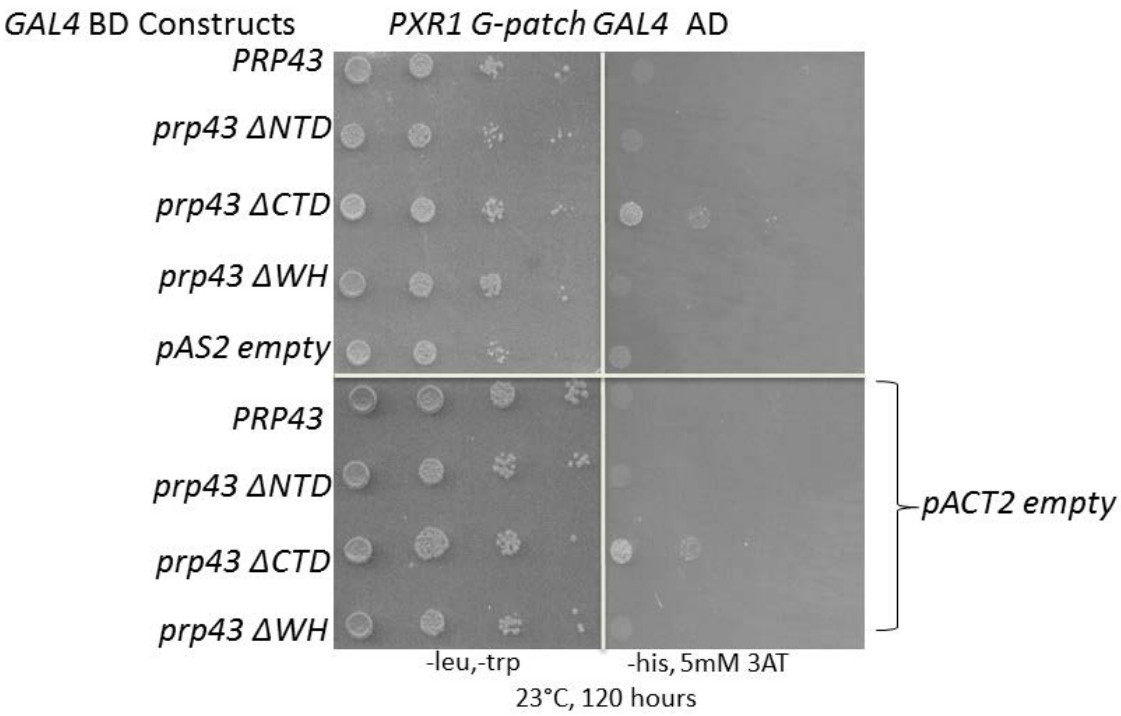

Figure 3.7.10. Removal of Prp43 domains fail to impact interaction with Pxr1 G-patch. Yeast-2 hybrid assay of Prp43 deletion derivatives and Sqs1 G-patch is shown as yeast growth assay. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel shows simple growth to select for yeast-2 hybrid plasmids, right panel shows reporter transactivation.

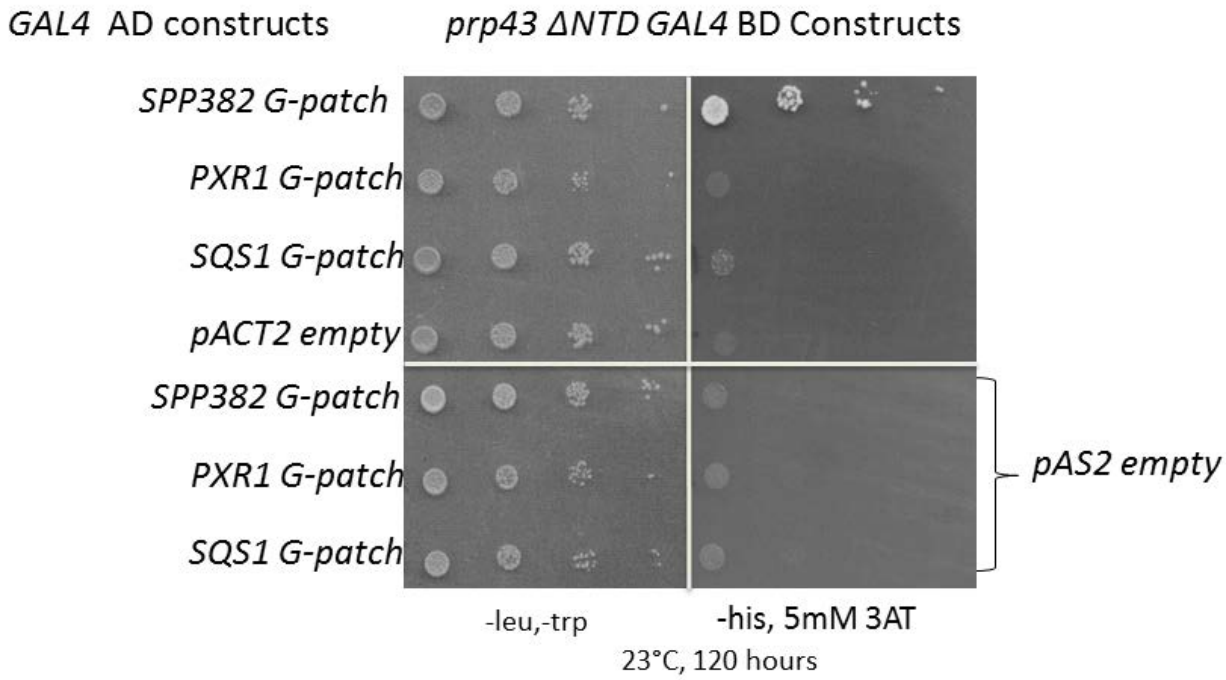

Figure 3.7.11. Gradient of Y2H activity between *prp43 ΔNTD* and isolated G-patch domains, such as Spp382 G-patch > Sqs1 G-patch >= Pxr1 G-patch. Yeast-2 hybrid assay of *prp43ΔNTD* and isolated G-patch domains are shown as yeast growth assay. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel shows simple growth to select for yeast-2 hybrid plasmids, right panel shows reporter transactivation.

pJ69-4a (pAS2-*prp43 ΔNTD* domain) with pACT2-	β-gal activity (Miller units) n=3	P value (Unpaired t-test) (Empty vs G-patches)
SPP382 G-patch	2.63±SD 0.16	<.0001*
PXR1 G-patch	1.27±SD .03	<.0001*
SQS1 G-patch	1.13±SD .01	<.0001*
Empty	0.67±SD .05	

Table 3.2. β-galactosidase activity of yeast strain pJ69-4a harboring the indicated pAS2-*prp43ΔNTD* domain and pACT2- isolated G-patch constructs. It is surprising that the activation of the reporter *GAL7-LacZ* gene shows only a 2-fold increase in the *prp43ΔNTD*-Spp382 G-patch pair compared to others and yet shows a significant difference in growth resulting from *GAL1-HIS3* reporter gene trans-activation (Figure 3.7.8 and 3.7.11).

3.7.3 Prp43 WH domain interacts with isolated G-patch fragment.

We next wanted to test if the Prp43 WH domain is sufficent to interact with the isolated G-patch domain utilizing the available yeast-2 hybrid contructs. Figure 3.7.12 shows interaction of isolated Spp382 G-patch with Prp43 WH domain. We observe a high level of autostimulation with the isolated WH domain, requiring the use of 50 mM 3-amino triazole to reduce background to an acceptable level. Although the NTD, CTD, RecA1, RecA2 or Ratchet showed no detectable activity, the WH domain shows activity above background. Extending this analysis with the isolated Pxr1 or Sqs1 G-patch segments yielded similar results (Figures 3.7.13 and 3.7.14).

Surprisingly, while Spp382 protein generally showed the best Y2H interaction with Prp43 in the earlier assays (Figure 3.1), when the isolated G-patch was scored against the isolated Prp43 WH domain, the the strongest activity was found for Pxr1 G-patch followed by Sqs1 G-patch and finally Spp382 (Figure 3.7.15 and Table 3.3). This result and the differential effect of the removal of the N-terminal region of Prp43 on G-patch domain binding suggest differences in how this common feature of Spp382, Pxr1 and Sqs1 may interact with Prp43.

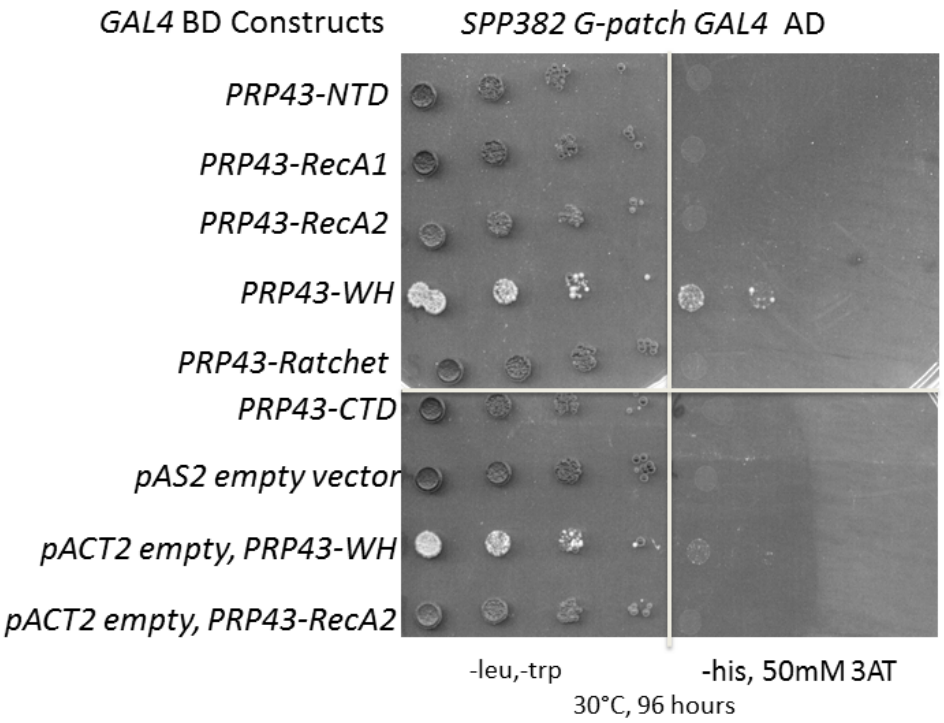

Figure 3.7.12. Prp43 WH domain interacts modestly with isolated Spp382 G-patch. Yeast-2 hybrid assay of Prp43 WH domain and Spp382 G-patch is shown as yeast growth assay. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel shows simple growth to select for yeast-2 hybrid plasmids, right panel shows reporter transactivation.

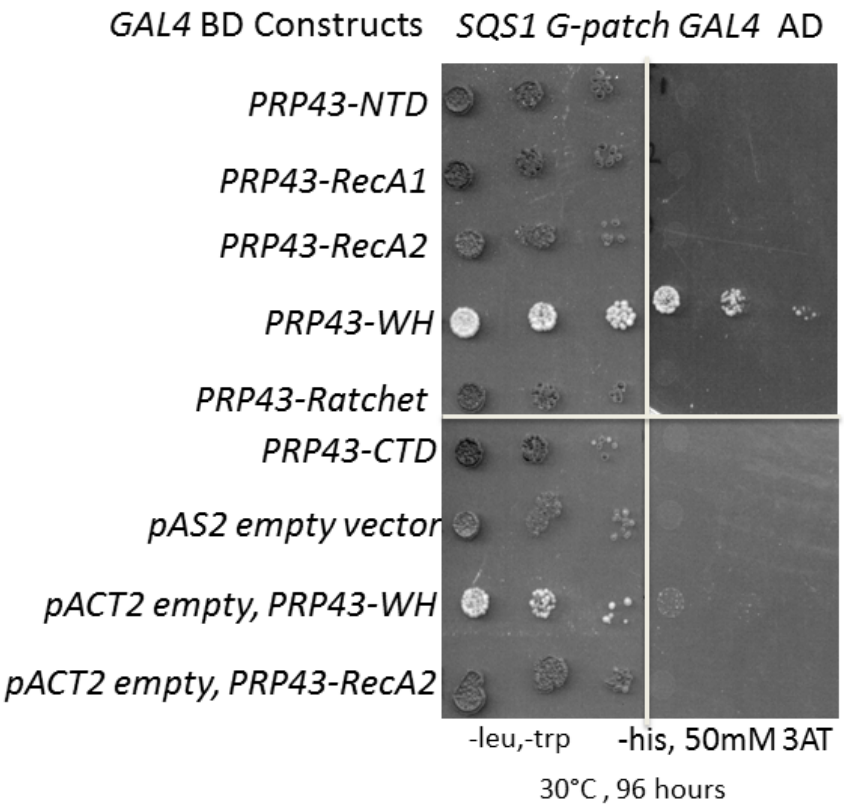

Figure 3.7.13. Prp43 WH domain interacts with isolated Sqs1 G-patch. Yeast-2 hybrid assay of Prp43 WH domain and Sqs1 G-patch is shown as yeast growth assay. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel shows simple growth to select for yeast-2 hybrid plasmids, right panel shows reporter transactivation.

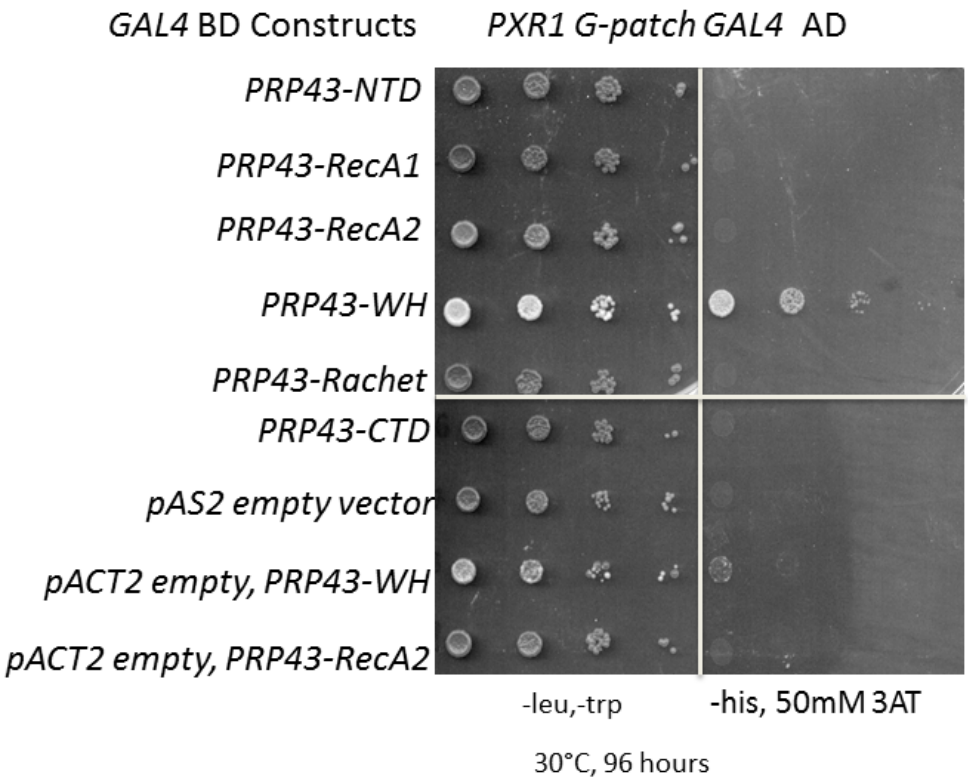

Figure 3.7.14. Prp43 WH domain interacts with isolated Pxr1 G-patch. Yeast-2 hybrid assay of Prp43 WH domain and Pxr1 G-patch is shown as yeast growth assay. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel shows simple growth to select for yeast-2 hybrid plasmids, right panel shows reporter transactivation.

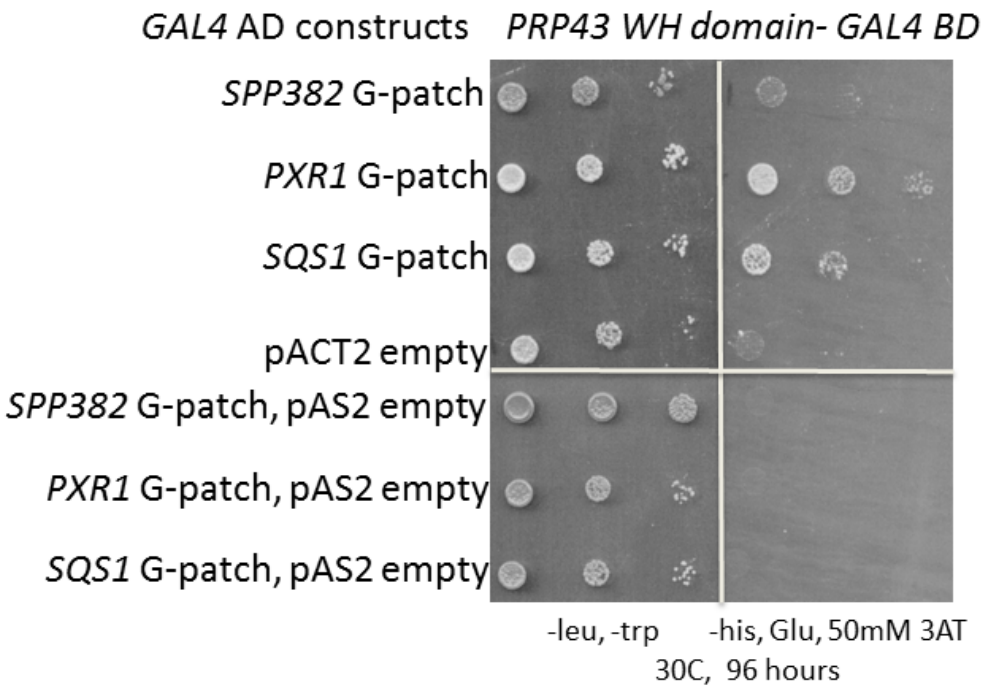

Figure 3.7.15. Gradient of Y2H activity between Prp43 WH domain and isolated G-patch domains, such as Pxr1>Sqs1>Spp382. Yeast-2 hybrid assay of Prp43 WH domain and G-patch domains are shown as yeast growth assay. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel shows simple growth to select for yeast-2 hybrid plasmids, right panel shows reporter transactivation.

pJ69-4a (pAS2-Prp43 WH domain) with pACT2-	β-Gal activity (Miller units) n=3	P value (Unpaired t-test) (Empty vs G-patches)
SPP382 G-patch	80.01±SD 3.74	.0018*
PXR1 G-patch	144.84±SD 2.71	<.0001*
SQS1 G-patch	127.21±SD 4.18	<.0001*
Empty	61.47±SD 2.24	

Table 3.3. β-galactosidase activity of yeast strain pJ69-4a harboring the indicated pAS2- Prp43 WH domain and pACT2- isolated G-patch constructs.

3.7.4 The Prp43 WH domain interacts with the Pxr1 102-149 aa segment

I previously showed that the Pxr1 102-149 region was necessary and sufficient for Ppr43 interaction (Figure 3.3.2 and 3.3.3). I next wanted to determine where on Prp43 this Pxr1 peptide might bind. To do this, I used the existing Y2H isolated Prp43 domain constructs paired with the isolated Pxr1 102-149 aa Figure 3.7. 16. Although the NTD, CTD, RecA1, RecA2 or Ratchet showed no detectable activity, the putative G-patch binding site, the WH domain, shows clear interaction by this assay.

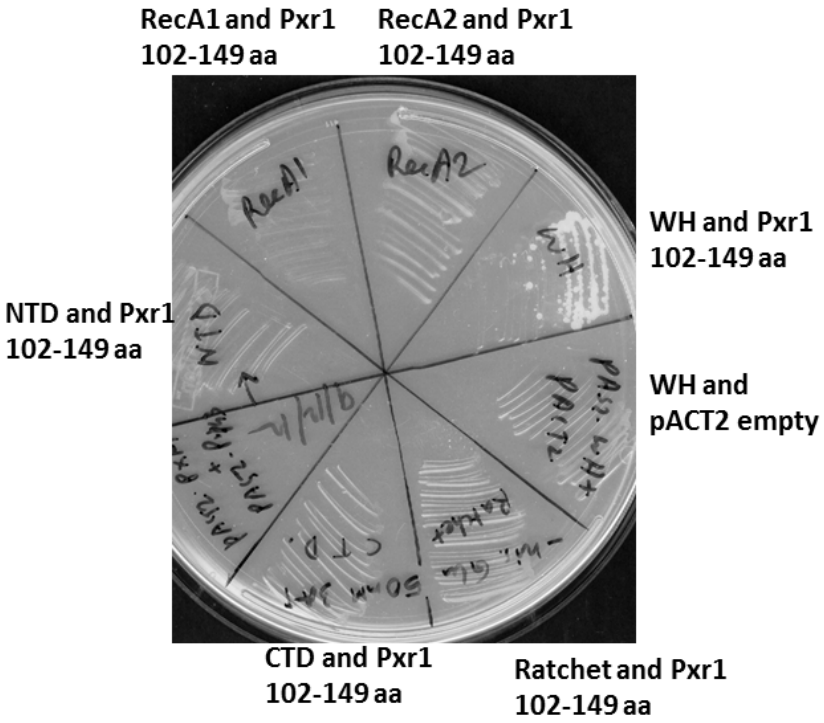

Figure 3.7.16. Prp43 WH domain interacts with isolated Pxr1 102-149 aa segment.

Yeast-2 hybrid assay of Prp43 domain constructs (pAS2-PRP43 NTD/ RECA1/RECA2/WH/RATCHET/CTD) or negative control and Pxr1 102-149 aa segment on a Y2H transactivation plate (-histidine, glucose plate with 50mM 3AT), incubated for 96 hours at 30°C. Note the control for WH autostimulation.

CHAPTER 4: DISCUSSION

In this study, I sought to understand the function of the G-patch domain using three *S. cerevisiae* G-patch domain bearing proteins that bind the DExD/H box- helicase Prp43, namely, Spp382, Pxr1 and Sqs1. Based on prior published and unpublished observations, I hypothesized that the G-patch domain within this protein set serves as an alternate Prp43 binding surface for recruitment of Prp43 to the pre-mRNA splicing (via Spp382) and rRNA processing (via Pxr1 & Sqs1) pathways. I tested this hypothesis, in addition to more generally probing the sites of interaction between Prp43 and the G-patch protein set.

The data presented in this study supports the view that the Spp382, Pxr1 and Sqs1 G-patch motifs can serve as Prp43 protein tethering sites and also provides evidence that G-patch sequence identity serves a cellular function more complex than simple Prp43 recruitment. Both the placement of the G-patch motif and reliance on the G-patch for Prp43 interaction and G-patch protein function differs among the proteins in this group. For instance, I localized the major Prp43 binding site of Pxr1 to a region well removed from the G-patch domain and showed that while loss of this high affinity Prp43 binding is tolerated, the loss of the Pxr1 G-patch is not. The results of peptide binding and mutagenesis studies suggest that the unanticipated possibility that while the Spp382, Pxr1 and Sq1 G-patch motifs are structurally similar, they may differentially interact with distinct regions of the Prp43 protein. Combined, the data support the view that the G-patch domain is more than a Prp43 recruitment feature but serves a function critical in pathway specific RNP substrate selection or Prp43 activation.

4.1 G-patch sequence identity impacts protein function in splicing or rRNA processing.

To gain insight into how the G-patch domain may contribute to Prp43's alternative roles in rRNA and pre-mRNA processing, I tested the modular nature of this element through domain swap experiments. This approach was previously used, for example, to establish the modular nature of the serine/arginine rich domain within the SR family of RNA processing factors (Chandler, Mayeda et al. 1997, van Der Houven Van Oordt, Newton et al. 2000). For instance, Fu and colleagues used domain swap along with splicing commitment assays to demonstrate that the arginine/serine rich domains of SC35and SF2/ASF are interchangeable while the RNA recognition motif (RRM) determines substrate specificity (Chandler, Mayeda et al. 1997). Likewise, a domain swap experiment between shuttling and non-shuttling SR proteins by Krainer and group identified the C-terminal RS domain of SF2/ASF as essential for nucleo-cytoplasmic shuttling (Caceres, Screaton et al. 1998).

Substitution of the Spp382 G-patch by the Sqs1 G-patch partially reconstitutes interaction with Prp43 in the Y2H assay (Figure 3.4.2) and Spp382 function in splicing (Figure 3.4.3). In contrast, the Pxr1 G-patch substitution into Spp382 fails in both assays. This difference in chimera function shows that at least some of the amino acid features which distinguish the three G-patch motifs are relevant to function. While Prp43 binding may be considered a shared activity for the Spp382, Pxr1 and Sqs1 proteins, there is not a 1:1 correspondence between Prp43 binding and the ability of the chimera to support biological function. For instance, while the Spp382- Pxr1 G-patch P48G chimera binds Prp43 while the Spp382- Pxr1 H55P G-patch chimera does not (or does so only very

poorly), but only the latter supports cell viability (Figure 3.6.1 and 3.6.2). On the other hand, the Spp382-Pxr1 G-patch D62M chimera binds Prp43 relatively well and also supports viability(Figure 3.6.1 and 3.6.2).

In this context of the Spp382- Pxr1 G-patch P48G chimera binding Prp43 but not supporting Spp382 activity, it is noteworthy that the Schwer group demonstrated that the lethal alanine or valine mutations in the Thr-384 position of the Prp43 RecA2 helicase domain produces a protein that binds Spp382 but its helicase activity is no longer enhanced by this G-patch protein (Tanaka, Aronova et al. 2007). If the Spp382 G-patch binds the RecA2 domain, this suggests that this association is required beyond Prp43 recruitment and may be directly relevant to Prp43's helicase contribution in lariat intron removal in pre-mRNA splicing (Tanaka, Aronova et al. 2007). How helicase stimulation is achieved by G-patch protein co-factor binding is unknown.

A possible trivial explanation for the failure of the Spp382-Pxr1 G-patch chimeric protein to support interaction with Prp43 or complement the *spp382::KAN* mutation is that this domain substitution renders the chimeric protein unstable or improperly folded. Unfortunately, there are no antibodies available against Spp382, Sqs1, Pxr1 or Prp43. I attempted western blot analysis with commercial Gal4 antibodies but could not reliably visualize the proteins produced by the Y2H constructs even when a strong Y2H signal was observed. However, while reduced protein stability cannot be ruled out, the smaller colony size of the Y2H strain pJ69-4a co-transformed with pAS2-Prp43 and pACT-*spp382*Δ*G-patch* with Pxr1 G-patch chimera (Figure 3.4.3, and Figure 3.6.1, left panel) suggests a modest dominant negative impact on growth, presumably dependent upon expression of this chimeric protein. Importantly, single amino acid substitutions in the

Pxr1 G-patch domain (R27G, H55P, K57E, D62M) in the Spp382-Pxr1 G-patch chimera render the chimeric protein functional (Figure 3.6.2). Also, I observe *in vitro* that the chimeric Spp382-Pxr1 G-patch and Spp382-Pxr1 H55P G-patch polypeptides bind Prp43 less well than the wildtype Spp382 peptide, reinforcing the fact that the chimeric protein, when stable, has reduced function (Figure A1).

It is noteworthy that pre-ribosomal processing is somewhat impaired in mutants that express certain Spp382-Pxr1 constructs (Figure 3.6.5). This defect correlates with poor mRNA splicing and ribosomal RNA processing efficiency of these strains (Figure 3.6.4). It is possible that adding a domain from an established pre-rRNA processing factor to a splicing factor recruits some of the Spp382 to the rRNA processing machinery in the nucleolus where it interferes with rRNA processing. However, it is also known that ribosomal proteins are required for rRNA processing, so disabling a general splicing factor like Spp382 may feed-back to impair ribosomal RNA processing efficiency, as the majority (~69%) of ribosomal protein encoding genes have introns (Ares, Grate et al. 1999, Spingola, Grate et al. 1999, Plocik and Guthrie 2012).

The reciprocal domain-swap experiment was performed with the *PXR1* gene and the results showed that the Sqs1 G-patch can substitute for the Pxr1 G-patch to support Pxr1function (Figure 3.4.5 and 3.4.6). Here again, this observation supports my initial hypothesis that the G-patch serves as a modular Prp43 binding surface. The inefficient pre-rRNA processing observed with the Pxr1- Spp382 G-patch chimera is consistent with the G-patch function beyond simple Prp43 binding and suggests that the G-patch may also bind pathway specific RNP features (RNA or protein) that differ between the pre-rRNP and pre-mRNP substrates. Conceivably, such a contribution to substrate selection

might promote the release of the lariat intron in splicing or a facilitate access or resolution of an rRNA intermediate or snoRNA/rRNA structure in rRNA processing, thus providing pathway specific substrate specificity.

Pxr1 is a component of the 90S pre-ribosomal processing particle and is involved in early pre-rRNA cleavage in site A1 and A2 in the 35S pre-rRNA (Guglielmi and Werner 2002, Lebaron, Papin et al. 2009). Sqs1, a component of the 90S, 60S and 40S pre-rRNA particles, is implicated in early 35S processing and also has a role in later stage of rRNA processing in 20S to 18S conversion (Lebaron, Papin et al. 2009, Pertschy, Schneider et al. 2009), see Figure 4.1. Supporting the G-patch's role in substrate selection, my data show that Spp382 G-patch substitution in Pxr1 causes a growth defect along with 35S pre-rRNA accumulation (Figure 3.4.6). In contrast, rRNA processing is quiet efficient with the Sqs1 G-patch substitution for the Pxr1 G-patch and does not cause a significant growth defect (Figure 3.4.5). It can be speculated that the Sqs1 G-patch substitution functions relatively efficiently because of possible similarity of the processing sites directed by Pxr1 and Sqs1 (Figure 4.1). Multiple potential Prp43 binding sites on the pre-rRNA were mapped by the Tollervey group (Bohnsack, Martin et al. 2009). Identification of the Prp43 binding site close to the 3' end of the 18S rRNA supports its function along with Sqs1 for 20S to 18S conversion, however, too little is known at this time about the mode of Prp43 binding and function to correlate these with specific sites of Pxr1 function. That the Pxr1-Spp382 G-patch chimera performs inefficiently may be because this domain interacts poorly with the local RNP structure (assuming that the G-patch interacts with RNA or associated protein) or because it appears to bind Prp43 differently. That is, the Spp382-G-patch binds the WH domain much less well than that

of Pxr1, yet has contacts elsewhere as revealed by the Prp43 NTD deletion and the Rec2 mutagenesis experiments of (Tanaka, Aronova et al. 2007). If the Spp382 G-patch binds elsewhere in Prp43, this might promote mis-alignment of the Pxr1-Spp382 chimeric protein on Prp43.

It is conceivable that I have mis-estimated the degree of RNA processing defects in mutant backgrounds by the Northern blot analyses due to how I normalized the samples. First, I used a fixed mass amount of extracted total cellular RNA based on OD_{260}, assuming that every yeast strain produces roughly equivalent amounts of RNA/cell. However, as at least 95% of the RNA reflected by the OD_{260} reading comes from the rRNA (Warner 1999), if rRNA production is greatly impaired, adjustment of sample recovery based on this absorbance reading could lead to an overloading of mRNA (relative to rRNA). For example a 50% reduction in rRNA in a sample due to experimental/genotype variation might lead to a twofold increase in the mRNA loaded (assuming that changes in rRNA abundance does not feedback to alter mRNA synthesis) . Fortunately, my use of the mRNA/pre-mRNA ratio will correct for differences in total RNA loaded when estimating splicing efficiency. However, as the rRNA intermediates were renormalized to the intronless *ADE3* mRNA, arbitrarily increased mRNA levels will result in an under-estimate the rRNA intermediate values.

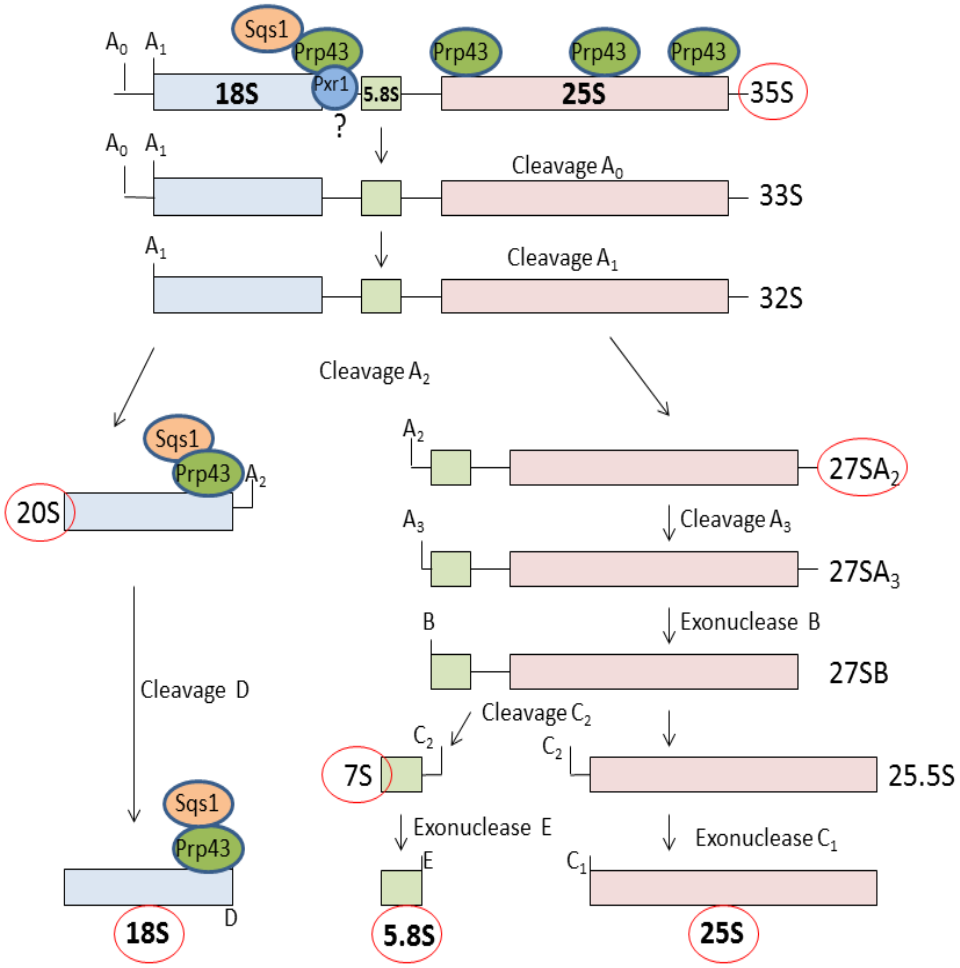

Figure 4.1. Schematic representation of the yeast rRNA processing pathway with Sqs1 and probable Pxr1 processing sites. The 35S precursor ribosomal RNA is processed by endonucleolytic and exonucleolytic cleavages into mature 25S, 18S and 5.8S rRNA. The intermediates steps studied by my northern blot are indicated by red circles. The Prp43 binding sites as identified by the Tollervey group (Bohnsack, Martin et al. 2009) are shown by green ovals. Genetic and biochemical evidences by the Henry and the Tollervey groups support the role of Sqs1 is assisting Prp43 for 20S pre-rRNA to 18S maturation (Lebaron, Papin et al. 2009, Pertschy, Schneider et al. 2009) (shown by orange oval).Although Pxr1 is directly associated with Prp43 in the 90S pre-rRNA

complex, and is involved in early pre-rRNA cleavage at A1 and A2 sites (Lebaron, Papin et al. 2009), the precise steps where Pxr1 assist Prp43 function is yet unknown (shown by blue oval and question mark).

4.2 The G-patch domains of Spp382, Pxr1 and Sqs1 are sufficient to bind Prp43 although they are not fully equivalent binding surfaces.

Using the yeast-2 hybrid approach, I showed that the G-patch domain is sufficient to interact with Prp43, although this interaction is comparatively weak by this assay when compared to the full length G-patch bearing proteins (Figure 3.1). This observation suggests that additional sites of Prp43 contact may exist within the G-patch proteins. At least for Pxr1 and Sqs1 this is clearly the case, as prior studies by our group (Pandit 2009) and others (Lebaron, Papin et al. 2009) showed that the N-terminus of Sqs1 contains a second Prp43 binding site and the work presented here identified a high-affinity Prp43 binding site within Pxr1 (amino acids 101-150). There are also differences in the G-patch dependence of the three proteins for Prp43 binding, with Spp382 interaction lost upon G-patch removal, Sqs1 interaction modestly reduced, and Pxr1 interaction not obviously diminished by removal of this motif (Figure 3.2).

The G-patch domain was proposed to be a RNA binding or a protein binding module(Aravind and Koonin 1999). Indeed, the Wolff group using fluorescence anisotropy with synthesized G-patch peptide (60 aa) and 7 bps RNA (U or A) homopolymers demonstrated the non-specific RNA binding activity of this domain from the *T. gondi* DNA repair enzyme (TgDRE) (Frenal, Callebaut et al. 2006). Leaving aside potential artifacts, our Y2H data provides the first evidence that the G-patch domain of Spp382, Sqs1 and Pxr1 are sufficient to function as a protein binding module. While exceptions are possible (e.g., two-hybrid interactions bridged by a third protein or nucleic acid factor *in vivo*), a positive Y2H interaction is generally considered reflective of direct protein-protein interaction (James, Halladay et al. 1996) and see review (Miller and

Stagljar 2004). This observation is consistent with earlier findings from the Lin group, the Cheng group, the Schwer group and the Henry group that suggested G-patch involvement in protein binding. Prior studies by the Lin group using Y2H assay and protein-protein binding assay showed that Prp2 C-terminal mutations (*W854A/L855A* or *D845N/C846Y*) lost interaction with the G-patch protein Spp2. The Y2H interaction was restored by a mutation of the conserved leucine in the Spp2 G-patch domain at amino acid position 109 to valine, although this mutation failed to restore function of the mutant Prp2 needed for cell viability. Using a dominant negative *prp2* allele (*prp2- D845L)*, the group showed the G-patch mutant Spp2 (L108V) modestly restored the splicing defect as well as interaction of the wildtype Spp2 and Prp2 (D845L), suggesting a possible role of the G-patch domain in DExD/H-box protein association (Silverman, Maeda et al. 2004). The Cheng group demonstrated that the 1-122 aa N-terminal Spp382 region containing the G-patch domain (61-108 aa) and some flanking sequence interacts with Prp43 in a Y2H assay (Tsai, Fu et al. 2005). This observation was corroborated by the Schwer group, which used a pull down assay to demonstrate that the Spp382 (1-120 aa) peptide binds Prp43 *in vitro* (Tanaka, Aronova et al. 2007). Likewise, the Henry group demonstrated that the C-terminal end of Sqs1 (574-767 aa) containing the G-patch (718-767 aa) plus considerable flanking sequence binds Prp43 in an *in vitro* pull-down assay (Lebaron, Papin et al. 2009). My study extends these earlier studies by showing that the G-patch is sufficient for Prp43 binding and in considering the role of family-specific G-patch features in pathway specific biological function.

Based on trans-activation of *GAL-HIS3* and *GAL7-LACZ* reporter genes, I find an apparent graded activity for Prp43 in the order Spp382>Sqs1>Pxr1 with both the full-

length proteins and the isolated G-patches. This differential binding is supported by a direct protein-protein pull down assay conducted with Spp382 N-terminal domain 121 amino acids containing either cognate Spp382 G-patch or heterologous G-patch from Pxr1/Sqs1 (Figure A1). Clearly there are lots of examples where differential ligand association modulates the degree of biological response. For example, the MacBeath group showed that high and low affinity epidermal growth factor receptor (EGFR) binding to ligand elucidate different signal transduction pathways in carcinoma cell lines (Krall, Beyer et al. 2011).In principle, the graded activity may be relevant to the Prp43 partitioning or function in rRNA processing and splicing through differential recruitment or activation of Prp43. This is an issue worthy of additional investigation.

4.3 Identification of the G-patch binding site within Prp43.

The crystal structure and domain organization of yeast Prp43 complexed with ADP-Mg^{2+} determined by the Henry group is provided below in Figure 4.2 from (Walbott, Mouffok et al. 2010). Based on modeling of the related Hel308-DNA bound structure, they propose that a single stranded nucleic acid binding cavity exists within Prp43 that is formed by the RecA1, RecA2, WH and the Ratchet domains. Based on this model, it is speculated that RecA1 binds to the 3' end of the single-stranded -nucleic acid and the RecA2 binds to the 5' end by binding the phosphate backbone. Analogous to the Hel308-DNA complex, the Prp43 ratchet helix has been proposed to act as a tether to hold the nucleic acid backbone in place so that the helicase can progress across the nucleic acid length. Although not tested by experimental validation, based on the Prp43 structure in complex with ADP-Mg^{2+}, the Henry group proposes that a model for Prp43 activity in which the ratchet domain pulls the RNA strand inside the cavity and the protruding β-hairpin structure from the RecA2 domain splits open the double stranded RNA strand (see review (Cordin, Hahn et al. 2012)).

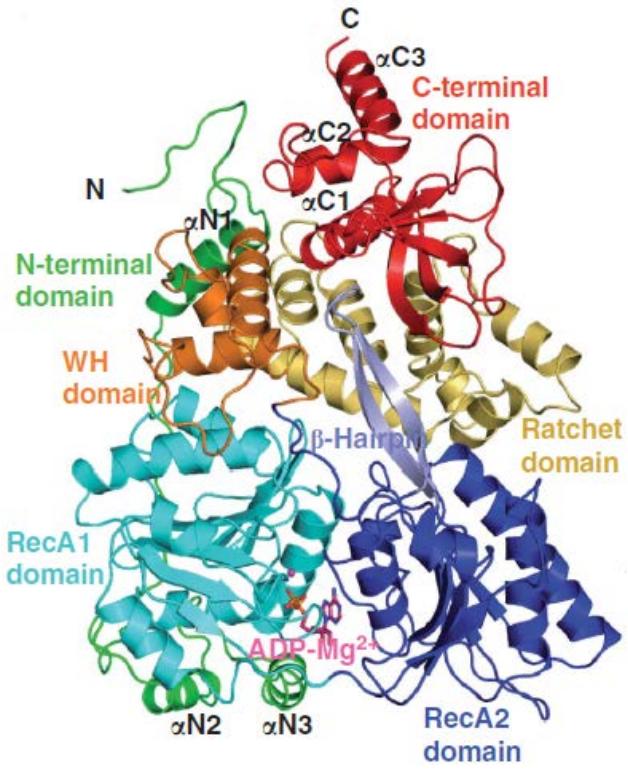

Figure 4.2. The structure of yeast Prp43 as identified by the Henry group shown in complex with ADP-Mg^{2+}. The six domains namely N-terminal domain (1-94 aa), RecA1 domain (95- 270 aa), RecA2 domain (271-457 aa), WH domain (458-521 aa), Ratchet domain (522- 635 aa) and C-terminal domain (636- 767 aa) are shown in different colors (Walbott, Mouffok et al. 2010).

To identify possible sites of G-patch interaction, I created series of Prp43 domain deletion mutants and single-domain insertion constructs to score by Y2H interaction(Figure 3.7.2-3.7.15).The selection of the deletion end points or the single domain insertions was guided by the domain structures defined by structural studies of the Henry and the Nielsen groups (He, Andersen et al. 2010, Walbott, Mouffok et al. 2010) (Figure 3.7.1).

My Y2H results show that the WH domain of Prp43 interacts with at least the Pxr1 and Sqs1 G-patch motifs and, in both cases, removal of the WH domain from Prp43 blocks this interaction. A simple interpretation of this observation is that the G-patch binds the WH domain and that this interaction is essential for both Pxr1 and Sqs1 to interact with the Prp43 DExD/H protein. However, at least the second point, that WH-domain/G-patch interaction is essential is not correct, as removal of the G-patch from Sqs1 or Pxr1 does not prevent the remainder of the protein from binding Prp43. The reliance on the WH domain but not its putative binding site (i.e., the G-patch) might reflect a distortion in Prp43 structure (to mask a second binding site) or protein instability when the WH-domain is removed. Protein instability appears not to be the cause, however, since the Spp382 protein continues to interact with the *prp43ΔWH* construct (Figure 3.7.2). This suggests that *prp43ΔWH* not only removes the G-patch binding site but distorts this protein in such a way to block Pxr1 and Sqs1 interactions in regions outside of the G-patch domains though contacts that are not critical for Spp382 association. Candidates for these contacts include the N-terminal region of Sqs1 (1-202 aa) (Lebaron, Papin et al. 2009)and the 101-150 amino acids region of Pxr1 both of which interact with the Prp43 protein (this study, Figure 3.3.2 and Figure 3.3.3).

Domain 4 of Prp43 is a winged helix (WH) domain. WH domains are found throughout nature in many types of proteins including such functionally relevant helicases as Hel308, Brr2 and Mtr4. This domain is suggested to function as both nucleic acid binding domain or protein binding domain although its role in helicase function is unknown, as reviewed in (Gajiwala and Burley 2000, Woodman and Bolt 2011). My results suggest that the WH domain may function as a G-patch binding module for Prp43. Since the WH domain appears to reside in the RNA binding cavity of Prp43, in principle, the G-patch association might serve to enhance or otherwise alter the putative WH/RNA substrate interaction by making stabilizing contacts to hold the RNA or protein in place or by inducing a conformational change within Prp43 to change its intrinsic RNA affinity. Alternatively (or in addition), the G-patch-WH domain association, if biologically relevant, may help orient the remainder of the Spp382, Sqs1 and Pxr1 G-patch protein within the RNP complex to promote favorable interactions with other regions of Prp43 or other protein or RNA features of the pre-rRNA or pre-mRNA splicing machineries.

Surprisingly, while the full-length Prp43 interacts with the G-patch proteins (or isolated G-patch domains) with apparent affinities Spp382>Sqs1>Pxr1 (Figure 3.1), the Y2H pattern is reversed when the isolated WH domain is scored for association. Here, the *GAL1-HIS3* and *GAL7-LACZ* reporter genes suggest the order Pxr1>Sqs1>Spp382 for the full-length G-patch proteins and the isolated G-patch domains with the isolated WH domain (Figure 3.7.13). Reciprocal Y2H assays with Gal4 activation domain fused to the Prp43 WH and the Gal4 binding domain fused to full length G-patch proteins show that the WH domain is sufficient to interact with Pxr1, but not quite as well with Spp382

and Sqs1 (Appendix Figure A.2.2). I acknowledge that the Spp382 G-patch interactions with the WH domain is extremely weak and only slightly above background (Figure 3.7.13). One possible explanation for this apparent discrepancy of isolated Prp43 WH domain versus full length Prp43 binding is that the three G-patch domains may not bind to the same site in Prp43. While the Pxr1 and Sqs1 G-patch possibly binds to the Prp43 WH domain, the Spp382 G-patch might bind to some other Prp43 region, possibly the RecA2 domain, which when mutated blocks Spp382 interaction (Tanaka, Aronova et al. 2007) and Figure 3.7.2 (this study). If this is the case, one might expect to find some level of sequence similarity between the Pxr1/Sqs1 and Spp382 G-patch binding sites in Prp43. Inspection of the RecA2 region reveals a 27 amino acid sequence with ~55% similarity with a portion of the WH domain. Intriguingly, the sequence of similarity includes the position that when mutated in the Prp43 construct, Y402A, reduced Spp382 binding *in vitro* (highlighted below) (Tanaka, Aronova et al. 2007).

```
RecA2  RKVVISTNIAETSLTIDGIVYV--VDP
       :. . :: .   .: :: .:.   .::
WH     RSNLSSTVLELKKLGIDDLVHFDFMDP
               10        20
```

Curiously, I observe that the full length Sqs1 and Pxr1 modestly interacts with isolated Prp43 RecA2 domain (Figure 3.7.6 and 3.7.7 respectively). This interaction might be reflective of this similarity between the WH domain and the RecA2 domain as shown above.

I found that the Prp43 Y2H interaction with all of the G-patch proteins was blocked by deletion of the RecA2 domain, whereas RecA1 removal blocked Spp382 and Sqs1 interaction, but not Pxr1 interaction, at least at 23°C (Figure 3.7.2, 3.7.3 and 3.7.4). This

suggests differences in G-patch protein binding where the intrinsically weakest interactor, Pxr1, is least dependent upon contacts presented (directly or indirectly) by the RecA1 domain. This reinforces my belief based largely on their otherwise unique structures that Pxr1, Spp32 and Sqs1 make distinct non-G-patch contacts on Prp43.

I observed that the removal of the N-terminal domain (NTD) from Prp43 does not inhibit its ability to interact with the full length G-patch proteins by the Y2H assay (Figure 3.7.2, 3.7 3, 3.7.4) indicating that this domain does not likely impair the stable association of Prp43 with the G-patch RNA processing proteins *in vivo*. Consistent with this interpretation, the Schwer group showed that the first 90 amino acid of Prp43 is dispensable for growth (Martin, Schneider et al. 2002), demonstrating that this domain is not required for Prp43's essential role in rRNA or pre-mRNA processing pathways. Interestingly, I find expression of the *prp43*ΔNTD protein partially ameliorates the cytotoxic effect of Sqs1 overexpression in the Y2H strain PJ69 (Figure 3.7.3) consistent with earlier observation by our laboratory that Sqs1 overexpression results in pre-mRNA splicing defects in otherwise wildtype yeast and that the associated growth inhibition can be partially suppressed with enhanced *PRP43* expression (Pandit 2009).

The N-terminus of Prp43 is not conserved in the other yeast DEAH/RHA helicases shown to be involved in pre-mRNA splicing (Prp2, Prp16, Prp22, Prp43) or rRNA processing (Dhr1, Dhr2). It is curious that I find deletion of the Prp43 N-terminus significantly enhances Spp382 G-patch domain interaction compared to what is observed with the full length Prp43 protein. This enhanced interaction is not observed when the Pxr1 or Sqs1 G-patches are assayed (Figure 3.7.8 and 3.7.11). The published structural

studies show that the NTD domain folds back over the Prp43 protein to make contacts with residues within the WH domain, the RecA1 domain, the Ratchet domain and the C-terminal region of Prp43, Figure 8 and (He, Andersen et al. 2010, Walbott, Mouffok et al. 2010). Thus, assuming the overall structure of the protein remains intact, removal of the NTD presumably alters access to multiple other domains for Spp382 G-patch interaction. The observation that the Spp382 G-patch interaction is enhanced in the *prp43ΔNTD* suggests that the NTD masks access of the Spp382 G-patch to its Prp43 target binding site. This might be the WH domain, the RecA1 domain, the Ratchet domain, the C-terminal region of Prp43 or other regions of Prp43 whose presentation changes when the NTD is removed. It is to be noted here that isolated Spp382 G-patch domain fails to interact with Prp43 deleted of each of the above mentioned domains (Appendix A3).

Prp43 is involved in the dissociation of the natural post-catalytic spliceosome (Arenas and Abelson 1997, Tsai, Fu et al. 2005) as well as in the discard of sub-optimal pre-mRNA/intermediates at the Prp16 and Prp22 dependent stages (Pandit, Lynn et al. 2006, Koodathingal, Novak et al. 2010, Mayas, Maita et al. 2010), and thus can be considered as a general disassembly factor, see review (Cordin, Hahn et al. 2012). Spp382 binds to the spliceosome prior to Prp43 association and is required by Prp43 to promote intron release. It is possible that the non-essential Prp43 N-terminal region may serve a regulatory function for this particular DEAH-box protein to prevent premature activation of the Prp43 enzyme by the G-patch interaction. In such a model, Spp382 G-patch independent interactions would recruit Prp43 to the spliceosome and conformational changes triggered during the splicing reaction might dissociate the NTD from the Prp43 helicase core to permit Spp382 G-patch access and activation of the Prp43 enzyme.

Alternatively, it is possible that the binding of the Spp382 G-patch to Prp43 facilitates the displacement of the NTD from the enzymatic core, making the enzymatic core accessible to the pathway specific RNA substrate. In either case, such models predict changes in the efficiency of recycling post-catalytic spliceosome in the absence of the NTD of Prp43. Furthermore, if such a model was correct and pre-mature activation of Prp43 occurs, one might predict that the *prp43ΔNTD* mutant would antagonize rather than suppress other spliceosome assembly mutants such as *prp38-1, prp8-1,* or *prp19-1* (Pandit, Lynn et al. 2006, Pandit 2009).

The Henry group used a pull down assay to show that interaction of a Sqs1 peptide consisting of the G-patch and flanking sequence (574-767 aa) is lost when the CTD is removed from Prp43 (Figure 6 in (Walbott, Mouffok et al. 2010)). Based on this negative result, they proposed that the Prp43 CTD binds the Sqs1 G-patch motif. As mentioned earlier, the Lin group showed that mutations in the Prp2 C-terminus that perturbed spliceosome binding also altered Spp2 interaction via the Spp2 G-patch domain (Silverman, Maeda et al. 2004). Investigation of the Prp43 and Prp2 C-terminal 100 amino acids using LALIGN software retrieved a 37 amino acid stretch with 78.4% similarity, as shown below. Interestingly, this segment includes the Prp2 residues (highlighted in yellow below) involved in spliceosome binding in the Lin groups study (Silverman, Maeda et al. 2004). The green is alanine residue that based on the crystal structure is in tight contact with the NTD of Prp43. If this is the binding site for Spp382, then removal of the NTD might enhance interaction.

```
     827                                          863
Prp2   SKYVLYQQLMLTSKEFIRDCLVIPKEEWLIDMVPQIF
       ...:.:...:::::.::    . . :::::..:  .
Prp43  AEWVIYNEFVLTSKNYIRTVTSV-RPEWLIEIAPAYY
     691                                        726
```

This suggests the Prp43 CTD might be involved in G-patch domain association. But, similar to the removal of the N-terminal domain, I find that the removal of the Prp43 C-terminal domain (CTD) also does not reduce the Y2H signals seen with the full length Spp382, Sqs1, and Pxr1 G-patch proteins (Figure 3.7.2, 3.7.3 and 3.7.4) indicating that this region of Prp43 is not critical for G-patch protein binding. It is to be noted here that this deletion removes 649-748 aa from the Prp43 CTD (635-767 aa, (Walbott, Mouffok et al. 2010)), thus retaining part of the Prp43 CTD, through which the G-patch proteins might continue to interact. However, as stated before, Pxr1 and Sqs1 do not depend upon a G-patch interaction for Prp43 binding, so this does not rule our association of these G-patch proteins through the C-terminus of Prp43. Spp382 is G-patch dependent for Prp43 binding, so either the Spp382 G-patch does not bind the Prp43 CTD or G-patch-independent Spp382 contacts are made outside the Prp43 CTD that stabilize this interaction.

The removal of the CTD from Prp43 does not change the Y2H signal when paired with the isolated G-patch peptides (Figure 3.7.9, 3.7.10 and & 3.7.11). However, as the overall interaction between the isolated G-patches and Prp43 is considerably weaker than that seen with full-length G-patch proteins (see Figure 3.1) the increased background (i.e., autostimulation) seen with the *prp43ΔCTD* construct makes interpretation of this observation problematic. Unlike with the WH-domain, I was not able to show interaction of the isolated Prp43 CTD with isolated G-patch peptide, suggesting that if the Prp43

CTD if it binds the G-patch that may not be sufficient for G-patch interaction. While sensitive to the fact that Y2H results can be mis-interpreted, it is also possible that the loss of G-patch peptide interaction noted in the pull-down study of the Henry group resulted from the artificial masking of the WH domain due to misfolding of the *prp43ΔCTD* protein. Therefore at this point, while I favor a model of G-patch-WH domain interaction, the WH domain, the RecA2 domain and the CTD regions of Prp43 remain candidates for other G-patch associations.

4.4 Identification of a novel Prp43 binding site within Pxr1.

The results presented here document a non-G-patch binding site for Prp43 within Pxr1 (Figure 3.3.2 and 3.3.3). The 271 aa Pxr1 protein has two annotated domains: the N-terminal glycine-rich 46 aa G-patch domain found in certain RNP complex proteins (Aravind and Koonin 1999) and, in its C-terminal half, a lysine rich KKE/D domain found in a number of nucleolar proteins involved with ribosome synthesis (e.g., Nop56, Nop58, Cbf5, Dbp3) (Gautier, Berges et al. 1997, Guglielmi and Werner 2002). I have shown that the G-patch domain of Pxr1 is dispensable for efficient Y2H interaction with Prp43 (Figure 2), thus demonstrating the existence of at least one other Prp43 binding interface. By deletion analysis, I identified a 48 amino acid peptide (102-149aa) as essential for Pxr1 interaction with Prp43 (Figures 3.3.2) and demonstrated that this peptide is sufficient for Prp43 binding (Figure 3.3.3).

Surprisingly, while the G-patch is essential, the major Prp43 binding site of Pxr1 (102-149aa) was found to be dispensable for Pxr1 function in cell growth and rRNA processing (Figure 3.3.4 and 3.3.6). Thus, while Pxr1 is thought to function principally as an activator of Prp43 enzyme activity, its strongest Prp43 binding site appears not to be required for function. At least two possible interpretations for this come to mind. First, it is possible that the key function of Pxr1 in rRNA processing occurs independent of Prp43 interaction so that when the 102- 149 aa segment is removed Prp43 association but not Pxr1 function is lost. Conceivably, Pxr1 may even act to stimulate another DExD/H box protein as 18 other DExD/H-box factors contribute to the rRNA processing pathway, see review (Kressler, Hurt et al. 2010). Alternatively, it is possible that the protein remaining in the *pxr1Δ102-149 aa* construct continues to interact with Prp43, albeit with reduced

efficiency through other weak contacts with Prp43 (e.g., the G-patch interaction) perhaps stabilized by contacts that Pxr1 makes with the RNA or additional protein components of the rRNA processing machinery.

Comparison of the 48 aa Prp43 binding peptide of Pxr1 with putative homologs of Pxr1 (Table 2.2 in Materials and methods) through the MUSCLE multiple sequence alignment program reveals that Prp43 –binding segment of Pxr1 is partially conserved (41-60% similarity within a stretch of 15-46 amino acids) in species from *C. elegans*, *D. melanogaster* to humans (Figure 4.3). This conservation in Pxr1 structure suggests conserved function in rRNA processing through Prp43 interaction. Although this Prp43 binding domain is not critical for function whether this domain subtly influences Prp43 activity in pre-rRNA processing remains open to further investigation.

I also observed that this 48 aa segment is capable of binding the isolated Prp43 WH domain in the Y2H assay (Figure 3.7.16). The generous interpretation of this observation is that binding of this segment to Prp43 WH domain is true and is biologically relevant. Alternatively, it is possible that this interaction is reflective of promiscuous Y2H activity of the WH domain, and needs to be cautiously interpreted along with additional controls (unrelated protein interaction with WH domain in the Y2H assay).

Figure 4.3. Multiple sequence alignment of the Prp43 binding sequence of Pxr1 across species using the MUSCLE program (Edgar 2004, Sonnhammer and Hollich 2005). The column colors which indicate the average BLOSUM62 score of: light blue≥3, dark blue≥1, light gray≥0.2, no color otherwise. Abbreviations: *Saccharomyces cerevisiae*- SC , *Schizosaccharomyces pombe*- SP, *Arabidopsis thaliana*- AT, *Drosophila melanogaster*- DM, *Caenorhabditis elegans*- CE, *Danio rerio*- DR, *Xenopus laevis*- XL, *Gallus gallus*- GG, *Mus musculus*- MM, *Bos Taurus*- BT and *Homo sapiens*- HS. These are the same Pxr1 sequences used for alignment earlier (Figure 3. 5).

4.5 Future directions:

Apart from providing a tethering surface for Prp43, my work provides evidence that the G-patch domains contribute additional information relevant to Prp43 function in the rRNA and pre-mRNA processing pathways. Although the three dimensional structure of Prp43 has been independently determined by the Henry and Nielsen groups, the mechanistic details of how the co-factors (Spp382, Sqs1 and Pxr1) activate the enzymatic activity of this DExD/H box helicase protein remain unknown.

4.5.1 Role of the Prp43 NTD to modulate Prp43 activity.

My work suggests that the non-essential N-terminal domain of Prp43 restricts access of the Spp382 G-patch to its binding site. As G-patch interaction is apparently required for Prp43 activation, this Prp43 N-terminal domain, in principle, may serve a function in regulating Prp43 activation. Clearly, such a function, if present, is not critical for Prp43 function as this N-terminal domain can be removed without loss of cell viability. At this time we do not know if this deletion alters the activation properties of the G-patch protein binding or changes the activity of Prp43 function, for instance, in splicing fidelity (Pandit, Lynn et al. 2006, Koodathingal, Novak et al. 2010, Mayas, Maita et al. 2010) or helicase activity (Tanaka, Aronova et al. 2007). Both of these activities might be addressed experimentally using the *prp43ΔNTD* mutant. Guided by results of the Schwer group (Tanaka, Aronova et al. 2007) and using a recombinant Prp43 clone generously provided by Dr. Beate Schwer, I have established an RNA stimulated helicase and ATPase assay in our laboratory. The differences in activation of a recombinant *prp43ΔNTD* protein and a wildtype Prp43 by the truncated Spp382 (1-121 aa) segment

can be tested using these assays. If the role of NTD is to act as a barrier for Spp382 G-patch accession to the helicase core, I expect to see an increased rate of helicase activity with the *prp43ΔNTD* protein compared to the wildtype Prp43. On the other hand, if the displacement of the NTD is critical for helicase activation, with the *prp43ΔNTD* protein, I expect to see promiscuous helicase activity, even in the absence of Spp382 activator.

An alternate way to address this question is to lock in the NTD interactions that occur between the NTD and the helicase, WH and CTD domains to learn if inhibiting the proposed displacement of the Prp43 NTD impairs function. This could be tested by artificially introducing disulfide bridges between the NTD and close contacts elsewhere in the protein (e.g., RecA1, the WH domain, the CTD) (He, Andersen et al. 2010, Walbott, Mouffok et al. 2010). These changes (with appropriate single-mutant controls) can be used to score for loss of Prp43 activity *in vivo* or *in vitro* using established assays for splicing (Boon, Auchynnikava et al. 2006), intron release (Martin, Schneider et al. 2002), G-patch protein binding (Tanaka, Aronova et al. 2007), and the suppression of spliceosome assembly defects (Pandit 2009). Using this approach, the Ann-Bjornsti group established a role of the N-terminal domain of yeast DNA topoisomerase in locking the Top1 clamp across the DNA duplex and identified conserved features of this enzyme (Palle, Pattarello et al. 2008). Since the natural cytoplasmic and nuclear environment reduces di-sulfide bonds, use of conditions that make the intracellular environment more oxidative might be necessary. Two exciting recent studies show that spontaneous and stable disulfide bonds can be made outside of the vesicular system or mitochondrial system in yeast by making the normally reducing cytoplasm environment more oxidative (Cumming, Andon et al. 2004, Ostergaard, Tachibana et al. 2004). Simple

growth in the presence of 1 mM oxidized glutathione after mutation of the *met17 (*previously named *met15)* and *glr1* genes (encoding O-acetyl homoserine-O-acetyl serine sulfhydrylase and glutathione oxidoreductase, respectively), results in a 45-fold increase in the intracellular ratio of oxidized to reduced glutathione and the recovery of >90% of target protein in the oxidized form (Ostergaard, Tachibana et al. 2004). If this model is correct, I expect the N-terminal locked Prp43 in the *glr1*mutant background to show temperature sensitivity or growth defects correlated with reduced Prp43 activity compared to the control strains.

4.5.2 Localizing and confirming the biological relevance of the G-patch interactions with 1) WH domain, 2) C-terminal domain, 3) RecA2 domain.

The WH domain of Prp43 has been annotated as "helicase associated domain" (Walbott, Mouffok et al. 2010), although, its precise cellular function is not clear. My study supports the view that the WH domain is a binding site for at least the Pxr1 and Sqs1 G-patch peptides. As a next step, these Y2H data should be corroborated using an independent methodology such as a protein-protein binding assay with recombinant proteins. Guided by the available Prp43 structures (He, Andersen et al. 2010, Walbott, Mouffok et al. 2010), mutations can be introduced in the Y2H constructs (pAS2-Prp43 WH) of this study to identify amino acid residues that lose interaction with the G-patch domain. Once these residues are identified in the WH domain, mutations can be introduced in the G-patch domain that restore Y2H interaction. The WH domain mutation could then be made in the context of full length Prp43 to study growth defect, temperature sensitivity, defects in lariat intron release or defects in pre-rRNA accumulation associated with this mutation *in vivo*. If such a defect is identified, this

mutation can be coupled with the G-patch domain mutation to see if the G-patch mutation that restored interaction is capable of suppressing the defect associated with the WH domain mutation. The Staley group showed that the dominant negative, cold sensitive *prp43Q423N* mutant is defective in pre-rRNA processing and accumulates 35S pre-rRNA along with a decrease in the downstream pre-rRNA intermediates (Leeds, Small et al. 2006). Also, the Schwer group showed that the lethal, dominant negative *prp43 T123A* mutant blocks release of the lariat intron from the spliceosome (Martin, Schneider et al. 2002). To address if Prp43 substrate binding is affected by a mutation in its G-patch interaction domain, Prp43 mutation can be introduced in these dominant negative *prp43* mutants to see if it relieves the temperature sensitivity or dominant negative phenotype. This sort of analysis will be stepping stone to address how the binding of the G-patch proteins alters Prp43 structure and modulates its pathway specific function.

APPENDICES

Appendix 1: Biochemical interaction of Spp382 (1-121aa) with Prp43 shows diminished binding of the chimeric Spp382 peptides.

To test the modular nature of the G-patch domain in *S. cerevisiae*, I did the domain swap experiment, swapping the Spp382 G-patch with either Pxr1 or Sqs1 G-patch and tested the resultant chimeric constructs for reconstitution of Spp382 function *in vivo* (Figure 3.4.3) and Prp43 interaction in the Y2H assay (Figure 3.4.2). The Y2H and *spp382::KAN* complementation assays showed that the Spp382-chimeric constructs have related but non-identical function. Furthermore, my mutational study of the Pxr1 G-patch suggests that Prp43 binding does not correlate tightly with the ability of the Spp382-Pxr1 chimera to function in splicing. However, the Y2H assay, while a common means to score for protein interaction, is clearly indirect in this measurement. To extend this analysis, I used a more direct protein binding assay to investigate Prp43-binding to Spp382 peptides bearing the Spp382, Pxr1, or Sqs1 G-patch domains.

To a first approximation, Prp43 is retained by the wildtype Spp382 peptide as well as the chimeric peptides at 50 mM NaCl concentration (lanes 7-10), compared to background control sample (lane 6). At higher stringency (200mM NaCl), Prp43 continues to bind the wildtype Spp382 peptide (lane 12), but its association with the Spp382-chimeric G-patch proteins with Pxr1 and Pxr1(H55P) G-patch is significantly compromised, while that of the Sqs1 chimera being approximately 50% greater (lane 13-15). The data is consistent with earlier findings by Schwer group that the first 121 amino acid of Spp382 is capable of binding Prp43 (Tanaka et al, 2007) and supports our Y2H assay result that the original Spp382 G-patch binds most strongly to Prp43 (Figure 3.4.2).

Given the subtle differences in binding with the set of Spp382-chimeras, however, no firm conclusion can be reached concerning the differential binding activity of Pxr1, Pxr1 (H55P) and Sqs1 G-patches.

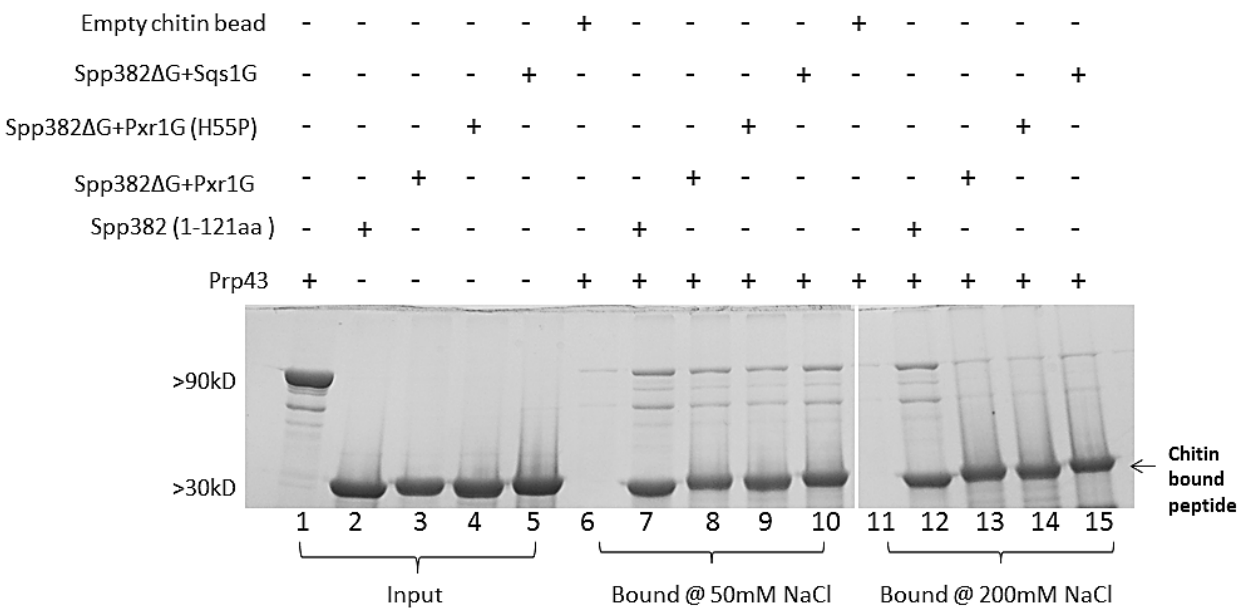

Figure A1. Differential binding of the Spp382-derived peptides to full-length Prp43. (Lanes 1-15 are all from one gel). The input Prp43 and the wildtype/chimeric Spp382 1-121 peptides are shown in lane 1-5.

Appendix 2: Reciprocal yeast2 hybrid assays

Reciprocal Y2H with Prp43 fused to the Gal4 activation domain (pACT2-Prp43) and full length G-patch proteins or isolated G-patch domains fused to the Gal4 DNA binding domain (pAS2-Spp382/Sqs1/Pxr1 or pAS2-Spp382 G patch/Sqs1 G patch/ Pxr1 G patch) corroborates earlier findings (Figure 1) that Prp43 interacts with full length Spp382, Sqs1 and Pxr1 or isolated G-patches in the Y2H assay such that Spp382>Sqs1>Pxr1.

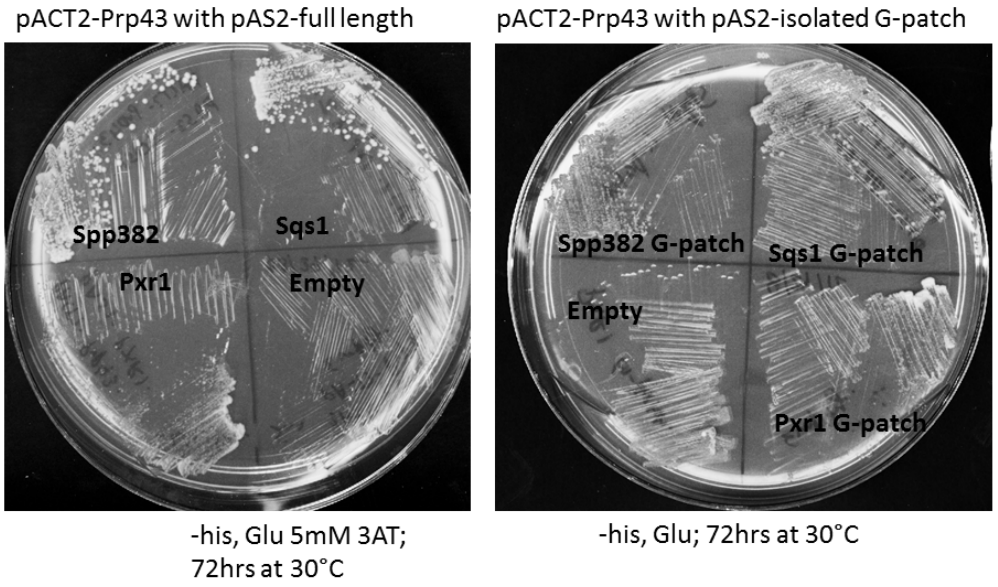

Figure A2.1. Reciprocal Y2H: Prp43 interacts with full length Spp382, Sqs1 and Pxr1 or isolated G-patches in the Y2H assay such that Spp382>Sqs1>Pxr1. Double transformants were streaked on transactivation plate and incubated for 72 hours at 30°C. NOTE: The isolated G-patches and Prp43 interaction is lost in presence of 5mM 3AT (data not shown).

Reciprocal Y2H experiments with Prp43 WH domain fused to the Gal4 activation domain (pACT2-Prp43 WH) and full length G-patch proteins fused to the Gal4 DNA binding domain (pAS2-Spp382/Sqs1/Pxr1) shows isolated Prp43 WH domain is sufficient to interact with Pxr1.

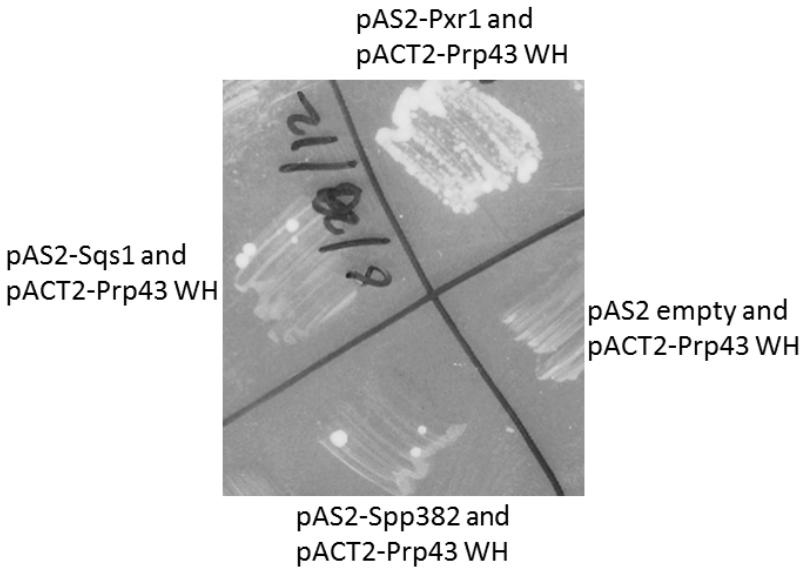

Figure A2.2. Reciprocal Y2H: Prp43 WH domain interacts with full length Pxr1. Double transformants were streaked on transactivation plates (-histidine, glucose media with 20mM 3AT) and incubated for 72 hours at 30°C.

Appendix 3: The isolated Spp382 G-patch fails to interact with *prp43ΔRecA1, prp43ΔRecA2, prp43ΔWH, prp43ΔRatchet* and *prp43ΔCTD domains*.

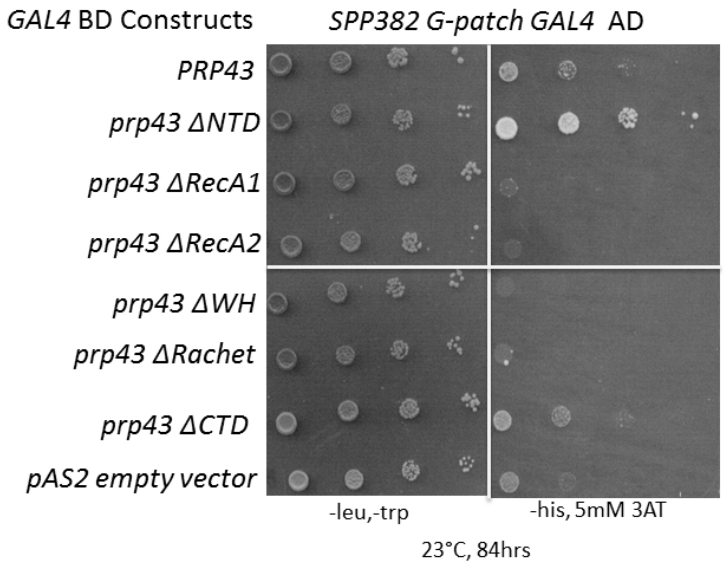

Figure A3.Yeast 2-hybrid interaction of Prp43 domain deletion derivatives and the isolated *SPP382 G-patch*. Each row represents isogenic strains, spotted as 10 serial dilutions and incubated under indicated conditions. Left panel represents simple growth to select for yeast-2 hybrid plasmids, right panel represents reporter transactivation. Note: *prp43ΔCTD*-Spp382 G-patch activity is almost comparable to *prp43ΔCTD* autostimulation.

This suggests that the isolated Spp382 G-patch is not sufficient to bind the Prp43 when the above domains are absent. One possibility is that one of these domains is a binding interface for the Spp382 G-patch. Alternatively, by the deletion of these domains, the protein conformation distorts in a way, making the Spp382 G-patch binding site inaccessible.

Appendix 4: Quantification of rRNA processing intermediates using Imagequant.

4.1 Quantification of rRNA processing intermediates of various Pxr1 deletion mutants from Figure 3.3.6

The quantification of rRNA processing intermediates or mature rRNAs were done using Imagequant software. The values are corrected for local background and normalized to that observed in the *PXR1* transformant.

Table A1: Fold change in rRNA precursors or mature rRNAs in the *PXR1* deletion mutants over wildtype *PXR1* in a *pxr1* null mutant background.

Vectors in *pxr1* null	35S	27SA2	23S	20S	25S	18S	7S	5.8S
Empty	2.4	1.5	1.5	.8	.8	1	.7	.9
PXR1	1	1	1	1	1	1	1	1
pxr1ΔG-patch	2.3	1.9	1.7	.8	.8	1	1.1	.9
pxr1Δ101-150	.6	1.2	.9	1.4	1	1.1	1.1	1.1
pxr1Δ101-226	2.3	1.5	1.2	.8	.5	.7	1.3	.5
pxr1Δ150-226	.5	.9	.9	1	1.1	1.1	1	1.2

4.2.1. Quantification of rRNA processing intermediates of chimeric Pxr1 from Figure 3.4.6

Table A2.1. Fold change in rRNA precursors or mature rRNAs in the chimeric *PXR1* over wildtype *PXR1* complemented *pxr1* null mutant. The values are corrected for local background and normalized to that observed in the *PXR1* transformant.

Vectors in *pxr1* null	35S	27SA2	23S	20S	25S	18S	7S	5.8S
Empty vector	6.2	2.5	2.5	.9	.9	1.2	.9	.9
PXR1	1	1	1	1	1	1	1	1
pxr1ΔG	4.1	1.7	2.6	.9	.5	.6	.6	.5
pxr1ΔG+PXR1 G	0.9	1	1.1	.9	.9	1	.9	1
pxr1ΔG+SPP382 G	2.6	.9	1.5	.5	.6	.7	.6	.6
pxr1ΔG+ SQS1 G	1.7	1.2	1.4	.8	.8	.9	.8	.9

4.2.2. Quantification of rRNA processing intermediates of chimeric Spp382 from Figure 3.4.6

Table A2.2. Fold change in rRNA precursors or mature rRNAs in the chimeric *SPP382* over *spp382ΔG+SPP382 G-patch* complemented N19 strain. The values are corrected for local background and normalized to that observed in the p*spp382ΔG+SPP382G* transformant.

Vectors in N19 strain	35S	27SA2	23S	20S	25S	18S	7S	5.8S
Empty vector	.5	.6	.8	1	1.1	1.2	1.2	1.1
spp382 ΔG	.8	.6	.9	.8	1	1	1	1
spp382	.7	1	.6	1.2	.9	1.1	1.2	1.1
spp382ΔG+SPP382 G	1	1	1	1	1	1	1	1
spp382ΔG+PXR1 G(H55P)	1	.8	1.3	1.2	1.4	1.4	1.3	1.3
spp382ΔG+ SQS1 G	.5	1.1	1.1	1.5	1.7	1.9	1.6	1.2

4.3 Quantification of varied ribosomal RNA processing efficiency in the Spp382-Pxr1 chimeras (Figure 3.6.5)

The quantification of rRNA processing intermediates or mature rRNAs were done using Imagequant software. The values are corrected for local background and normalized to that observed in the *spp382ΔG+Spp382G* transformant.

Table A3: Fold change in rRNA precursors or mature rRNAs in the *spp382* null mutant bearing the chimeric Spp382 with mutant Pxr1 G-patches.

Vectors in *spp382* null	35S	27SA2	23S	20S	25S	18S	7S	5.8S
SPP382	.8	1.2	1	1.3	1	1.2	1	1.2
spp382ΔG+SPP382G	1	1	1	1	1	1	1	1
spp382ΔG+pxr1G (R27G)	1.9	.6	2.2	1.3	2	2.2	2.4	1.9
spp382ΔG+pxr1G (H55P)	1.5	.7	2	1.3	1.6	1.6	1.9	1.3
spp382ΔG+pxr1G (K57E)	3	.8	3.2	1.5	1.5	1.6	1.2	.7
spp382ΔG+pxr1G(D62M)	2.9	.7	2.8	1.1	1.4	1.4	1.6	.8

Appendix 5: Yeast Strains used in this dissertation

Table A4: Yeast strains used in this dissertation.

	Strain	Figure number	Box, location
1	pAS2-Prp43+ pACT2-SPP382 G-patch only in pJ69-4a 1/19/12	3.1	40, 23
2	pAS2-Prp43+ pACT2-SQS1 G-patch only in pJ69-4a 1/19/12	3.1	40, 24
3	pAS2-Prp43+ pACT2-PXR1 G-patch only in pJ69-4a 1/19/12	3.1	40, 25
4	pAS2-PRP43 and pACT-SPP382 in PJ69-4a on 08/31/12 DB	3.1	41, 1
5	pACT-SQS1 and pAS2-Prp43 (12/20/11) DB	3.1	37, 31
6	pACT2-PXR1 and pAS2-Prp43 (12/12/11) DB	3.1	37, 23
7	pAS2-PRP43 and pACT empty in PJ69-4a on 08/31/12 DB	3.1	41, 6
8	pAS2-PRP43 and pACT-SPP382ΔG in PJ69-4a on 08/31/12	3.2	41, 2
9	pAS2-Prp43 and pACT2-PXR1ΔG-patch in PJ69 11/11/11	3.2	36, 57
10	PJ69-4A pAS2-PRP43 pACT-sqs1ΔG6 (#1), frozen 12/05/08 (SH)	3.2	33, 12
11	pAS2-Prp43 and pACT2-PXR1Δ7-101aa in PJ69 11/11/11	3.3.2	36, 58
12	pAS2-Prp43 and pACT2-PXR1Δ101-150aa in PJ69 11/11/11	3.3.2	36, 59
13	pAS2-Prp43 and pACT2-PXR1Δ25-150aa in PJ69 11/11/11	3.3.2	36, 60
14	pAS2-Prp43 and pACT2-PXR1Δ101-226aa in PJ69 11/11/11	3.3.2	36, 61
15	pAS2-Prp43 and pACT2-PXR1Δ150-226aa in PJ69 11/11/11	3.3.2	36,62
16	pAS2-Prp43 and pACT2-PXR1Δ7-150aa in PJ69 11/11/11	3.3.2	36, 63
17	pAS2-Prp43 and pACT2-PXR1Δ150-265aa in PJ69 11/11/11	3.3.2	36, 64

18	pACT2-PXR1(101-150aa) + pAS2-Prp43 in pJ69-4a(06/14/12)	3.3.3	39, 70
19	pAS2-Prp43 and pACT2-PXR1 (101-226aa) fragment in pJ69-4a (06/28/12)	3.3.3	40, 12
			37, 24
20	Ycplac111-PXR1 in Δpxr1:KAN (12/15/11) DB	3.3.4	37, 25
21	Ycplac111-PXR1ΔG-patch in Δpxr1:KAN (12/15/11) DB	3.3.4	37, 26
22	Ycplac111-PXR1Δ101-150aa in Δpxr1:KAN (12/15/11) DB	3.3.4	37, 27
23	Ycplac111-PXR1Δ101-226aa in Δpxr1:KAN (12/15/11) DB	3.3.4	37, 28
24	Ycplac111-PXR1Δ150-226aa in Δpxr1:KAN (12/15/11) DB	3.3.4	37, 29
25	Ycplac111 empty in Δpxr1:KAN (12/15/11) DB	3.3.4	37, 30
26	pAS2-PRP43 and pACT-SPP382ΔG+SPP382G in PJ69-4a on 05/13/10 DB	3.4.2	35, 26
27	pAS2-PRP43 and pACT-SPP382ΔG+PXR1G in PJ69-4a 08/31/12 DB	3.4.2	41, 3
28	pAS2-PRP43 and pACT-SPP382ΔG+SQS1G in PJ69-4a on 08/31/12 DB	3.4.2	41, 5
29	N19+Ycplac111-Spp382 12/12/11	3.4.3, 3.4.4B	37, 9
30	N19+Ycplac111-Spp382delG-patch 12/12/11	3.4.3, 3.4.4B	37, 10
31	N19+Ycplac111-Spp382delG-patch +Spp382 G-patch 12/12/11	3.4.3, 3.4.4B	37, 11
32	N19+Ycplac111-Spp382delG-patch +PXR1 G-patch 12/12/11	3.4.3, 3.4.4B	37, 12
33	N19+Ycplac111-Spp382delG-patch +SQS1 G-patch 12/14/11	3.4.3	37, 15
34	N19+Ycplac111 empty vector 12/12/11	3.4.3, 3.4.4B	37, 14
35	N19+Ycplac111-Spp382 FoA+ 09/19/12	3.4.4A, 3.4.6	41, 35
36	N19+Ycplac111-Spp382delG+Spp382 G	3.4.4A,	41, 36

	FoA+	3.4.6	
37	N19+Ycplac111-Spp382delG+Sqs1 G FoA+ 09/19/12	3.4.4A, 3.4.6	41, 37
38	pxr1:: KAN		28, 32
39	pxr1::KAN (Ycplac111-PXR1)	3.4.5, 3.4.6	35, 67
40	pxr1::KAN (Ycplac111-PXR1del G-patch)	3.4.5, 3.4.6	35, 68
41	pxr1::KAN (Ycplac111-PXR1del G-patch + Pxr1 G-patch)	3.4.5, 3.4.6	35, 69
42	pxr1::KAN (Ycplac111-PXR1del G-patch + Sqs1 G-patch)	3.4.5, 3.4.6	35, 76
43	pxr1::KAN (Ycplac111-PXR1del G-patch + Spp382 G-patch)	3.4.5, 3.4.6	35, 75
44	pxr1::KAN (Ycplac111-PXR1del G-patch + Pxr1 G-patch H55P)	3.4.5, 3.4.6	35, 74
45	pxr1::KAN (Ycplac111 empty)	3.4.5, 3.4.6	35, 77
46	pAS2-Prp43 and pACT-Spp382delG+Pxr1G R27G in pJ69-4a	3.6.1	41, 7
47	pAS2-Prp43 and pACT-Spp382delG+Pxr1G P48G in pJ69-4a	3.6.1	41, 8
48	pAS2-Prp43 and pACT-Spp382delG+Pxr1G K57E in pJ69-4a	3.6.1	41, 9
49	pAS2-Prp43 and pACT-Spp382delG+Pxr1G D62M in pJ69-4a	3.6.1	41, 10
50	pAS2-Prp43 and pACT-SPP382delG+PXR1G H55P in PJ694a	3.6.1	41, 4
51	N19		12, 42
52	N19+ Ycplac111-Spp382delG+Pxr1G (R27G)	3.6.2	40, 71
53	N19+ Ycplac111-Spp382delG+Pxr1G (P48G)	3.6.2	40, 72
54	N19+ Ycplac111-Spp382delG+Pxr1G (K57E)	3.6.2	40, 73
55	N19+ Ycplac111-Spp382delG+Pxr1G (D62M)	3.6.2	40, 74
56	N19+Ycplac111-SPP382ΔG + PXR1G H55P	3.6.2	35, 12
57	N19+Ycplac111-Spp382delG+Pxr1G (R27G) FoA+	3.6.3, 3.6.4,	41, 38

		3.6.5	
58	N19+Ycplac111-Spp382delG+Pxr1G (K57E) FoA+	3.6.3, 3.6.4, 3.6.5	41, 39
59	N19+Ycplac111-Spp382delG+Pxr1G (H55P) FoA+	3.6.3, 3.6.4, 3.6.5	41, 40
60	N19+Ycplac111-Spp382delG+Pxr1G (D62M) FoA+	3.6.3, 3.6.4, 3.6.5	41, 41
61	pACT-Spp382 and pAS2-Prp43	3.7.2	37, 1
62	pACT-Spp382 and pAS2-empty vector	3.7.2	37, 2
63	pACT-Spp382 and pAS2-Prp43del NTD #3	3.7.2	37, 3
64	pACT-Spp382 and pAS2-Prp43del RecA1 #20	3.7.2	37, 4
65	pACT-Spp382 and pAS2-Prp43 delRecA2 # 4	3.7.2	37, 5
66	pACT-Spp382 and pAS2-Prp43 delWH #17	3.7.2	37, 6
67	pACT-Spp382 and pAS2-Prp43 del Rachet #12	3.7.2	37, 7
68	pACT-Spp382 and pAS2-Prp43 del CTD #5	3.7.2	37, 8
69	pACT2-PXR1 and pAS2-Prp43del NTD #3	3.7.4	37, 16
70	pACT2-PXR1 and pAS2-Prp43del RecA1 #20	3.7.4	37, 17
71	pACT2-PXR1 and pAS2-Prp43 delRecA2 # 4	3.7.4	37, 18
72	pACT2-PXR1 and pAS2-Prp43 delWH #17	3.7.4	37, 19
73	pACT2-PXR1 and pAS2-Prp43 del Rachet #12	3.7.4	37, 20
74	pACT2-PXR1 and pAS2-Prp43 del CTD #5	3.7.4	37, 21
75	pACT2-PXR1 and pAS2-empty vector	3.7.4	37, 22
76	pACT2-PXR1 and pAS2-Prp43	3.7.4	37, 23
77	pACT-SQS1 and pAS2-Prp43	3.7.3	37, 31
78	pACT- SQS1 and pAS2-empty vector	3.7.3	37, 32
79	pACT- SQS1 and pAS2-Prp43del NTD #3	3.7.3	37, 33
80	pACT- SQS1 and pAS2-Prp43del RecA1	3.7.3	37, 34

	#20		
81	pACT- SQS1 and pAS2-Prp43 delRecA2 # 4	3.7.3	37, 35
82	pACT- SQS1 and pAS2-Prp43 delWH #17	3.7.3	37, 36
83	pACT- SQS1 and pAS2-Prp43 del Rachet #12	3.7.3	37, 37
84	pACT- SQS1 and pAS2-Prp43 del CTD #5	3.7.3	37, 38
85	pACT-Spp382 + pAS2-Prp43 NTD fragment in PJ69-4a	3.7.5	39, 17
86	pACT-Spp382 + pAS2-Prp43 RecA1 fragment in PJ69-4a	3.7.5	39, 18
87	pACT-Spp382 + pAS2-Prp43 RecA2 fragment in PJ69-4a	3.7.5	39, 19
88	pACT-Spp382 + pAS2-Prp43 WH fragment in PJ69-4a	3.7.5	39, 20
89	pACT-Spp382 + pAS2-Prp43 Ratchet fragment in PJ69-4a	3.7.5	39, 21
90	pACT-Spp382 + pAS2-Prp43 CTD in PJ69-4a	3.7.5	39, 22
91	pACT-SQS1 + pAS2-Prp43 NTD fragment in PJ69-4a	3.7.7	39, 36
92	pACT- SQS1 + pAS2-Prp43 RecA1 fragment in PJ69-4a	3.7.7	39, 37
93	pACT- SQS1 + pAS2-Prp43 RecA2 fragment in PJ69-4a	3.7.7	39, 38
94	pACT- SQS1 + pAS2-Prp43 WH fragment in PJ69-4a	3.7.7	39, 39
95	pACT- SQS1 + pAS2-Prp43 Ratchet fragment in PJ69-4a	3.7.7	39, 40
96	pACT- SQS1 + pAS2-Prp43 CTD in PJ69-4a	3.7.7	39, 41
97	pACT2-SPP382 G-patch & pAS2 Prp43delNTD	3.7.8	38, 43
98	pACT2-SPP382 G-patch & pAS2 Prp43del RecA1	3.7.8	38, 44
99	pACT2-SPP382 G-patch & pAS2 Prp43del RecA2	3.7.8	38, 45
100	pACT2-SPP382 G-patch & pAS2 Prp43del WH	3.7.8	38, 46
101	pACT2-SPP382 G-patch & pAS2 Prp43del Ratchet	3.7.8	38, 47
102	pACT2-SPP382 G-patch & pAS2 Prp43del	3.7.8	38, 48

	CTD		
103	pACT2-Pxr1 G-patch only and pAS2-Prp43del NTD in pJ69	3.7.10	39, 63
104	pACT2-Pxr1 G-patch and pAS2-Prp43delWH in pJ69-4a	3.7.10	39, 54
105	pACT2-Pxr1 G-patch only and pAS2-Prp43 in pJ69-4a	3.7.10	39, 61
106	pACT2-Pxr1 G-patch only and pAS2-Prp43del CTD in pJ69	3.7.10	39, 65
107	pACT2-Sqs1 G-patch and pAS2-Prp43delWH in pJ69-4a	3.7.9	39, 53
108	pACT2-Sqs1 G-patch only and pAS2-Prp43 in pJ69-4a	3.7.9	39, 60
109	pACT2-Sqs1 G-patch only and pAS2-Prp43del NTD in pJ69	3.7.9	39, 62
110	pACT2-Sqs1 G-patch only and pAS2-Prp43del CTD in pJ69	3.7.9	39, 64
111	pACT2-Spp382 G-patch only + PAS2 empty in PJ69-4a	3.7.12	38, 74
112	pACT2-Spp382 G-patch only + PAS2 Prp43 NTD in PJ69-4a	3.7.12	38, 75
113	pACT2-Spp382 G-patch only + PAS2 – Prp43 RecA1 in PJ69	3.7.12	38, 76
114	pACT2-Spp382 G-patch only + PAS2 – Prp43 RecA2 in PJ69	3.7.12	38, 77
115	pACT2-Spp382 G-patch only + PAS2 – Prp43 RecA1+ReCA2	3.7.12	38, 78
116	pACT2-Spp382 G-patch only + PAS2 – Prp43 WH in PJ69-4a	3.7.12	38, 79
117	pACT2-Spp382 G-patch only + PAS2 – Prp43 Ratchet in PJ69-4a	3.7.12	38, 80
118	pACT2-Spp382 G-patch only + PAS2 – Prp43 CTD in PJ69-4a	3.7.12	38, 81
119	pACT2-SQS1 G-patch only + pAS2-Prp43 NTD fragment in PJ69-4a	3.7.13	39, 24
120	pACT2-SQS1 G-patch only + pAS2-Prp43 RecA1 in PJ69-4a	3.7.13	39, 25
121	pACT2-SQS1 G-patch only + pAS2-Prp43 RecA2 in PJ69-4a	3.7.13	39, 26
122	pACT2-SQS1 G-patch only + pAS2-Prp43 WH fragment in PJ69-4a	3.7.13	39, 27

123	pACT2-SQS1 G-patch only + pAS2-Prp43 Ratchet in PJ69-4a	3.7.13	39, 28
124	pACT2-SQS1 G-patch only + pAS2-Prp43 CTD in PJ69-4a	3.7.13	39, 29
125	pACT2-PXR1 G-patch only + pAS2-Prp43 NTD fragment in PJ69-4a	3.7.14	39, 30
126	pACT2- PXR1 G-patch only + pAS2-Prp43 RecA1 in PJ69-4a	3.7.14	39, 31
127	pACT2- PXR1 G-patch only + pAS2-Prp43 RecA2 in PJ69-4a	3.7.14	39, 32
128	pACT2- PXR1G-patch only + pAS2-Prp43 WH fragment in PJ69-4a	3.7.14	39, 33
129	pACT2- PXR1 G-patch only + pAS2-Prp43 Ratchet in PJ69-4a	3.7.14	39, 34
130	pACT2- PXR1 G-patch only + pAS2-Prp43 CTD in PJ69-4a	3.7.14	39, 35
131	pACT2-PXR1(101-150aa) + pAS2-Prp43 in pJ69-4a	3.7.16	39, 70
132	pACT2-PXR1(101-150aa) + pAS2-Prp43 NTD fragment in pJ69-4a	3.7.16	39, 71
133	pACT2-PXR1(101-150aa) + pAS2-Prp43 RecA1 in pJ69-4a	3.7.16	39, 72
134	pACT2-PXR1(101-150aa) + pAS2-Prp43 RecA2 in pJ69-4a	3.7.16	39, 73
135	pACT2-PXR1(101-150aa) + pAS2-Prp43 WH fragment in pJ69-4a	3.7.16	39, 75
136	pACT2-PXR1(101-150aa) + pAS2-Prp43 Ratchet in pJ69-4a	3.7.16	39, 76
137	pACT2-PXR1(101-150aa) + pAS2-Prp43 CTD fragment in pJ69-4a	3.7.16	39, 77
138	pACT2-Prp43 WH domain and pAS2 empty in pJ69-4a 09/05/12	3.7.16	41, 23
139	pACT2 empty + pAS2-Prp43 del CTD in pJ69-4a	3.7.8, 3.7.9, 3.7.10	40, 7
140	pACT2 empty+ pAS2-Prp43 RecA2 in pJ69-4a	3.7.12, 13, 14	39, 44
141	pACT2-Prp43 and pAS2-Spp382 in pJ69-4a	A2.1	41, 13
142	pACT2-Prp43 and pAS2-Pxr1 in pJ69-4a	A2.1	41, 14
143	pACT2-Prp43 and pAS2-Sqs1 in pJ69-4a	A2.1	41, 15
144	pACT2-Prp43 and pAS2 empty in pJ69-4a	A2.1	41, 19

145	pACT2-Prp43 and pAS2-Spp382 G-patch in pJ69-4a	A2.1	41, 16
146	pACT2-Prp43 and pAS2-Pxr1 G-patch in pJ69-4a	A2.1	41, 17
147	pACT2-Prp43 and pAS2-Sqs1 G-patch in pJ69-4a	A2.1	41, 18
148	pACT2-Prp43 WH domain and pAS2-Spp382 in pJ69-4a	A2.2	41, 20
149	pACT2-Prp43 WH domain and pAS2-PXR1 in pJ69-4a	A2.2	41, 21
150	pACT2-Prp43 WH domain and pAS2-Sqs1 in pJ69-4a	A2.2	41, 22
151	pACT2-Prp43 WH domain and pAS2 empty in pJ69-4a	A2.2	41, 23

Appendix 6: *E.coli* strains containing plasmids used in this dissertation

Table A5: *E.coli* strains containing plasmids used in this dissertation

	Strain name	Box location
1	pAS2-Prp43 in TG1	16, 63
2	pACT2-SPP382 G-patch only #16	15, 30
3	pACT2-PXR1 G-patch only #12	16, 3
4	pACT2-SQS1 G-patch only #1	16, 4
5	pACT-SPP382	12. 34
6	pACT2-PXR1 clone #1 in TG1	15, 61
7	PACT-Sqs1	12, 32
8	pACT2 empty vector in TG1	15, 22
9	pACT2-PXR1ΔG clone #2 in TG1	15, 62
10	pACT-SPP382 ΔG in TG1	14, 59
11	pACT2-PXR1 del 7-101 in TG1	16, 64
12	pACT2-PXR1 del 25-150 in TG1	16, 65
13	pACT2-PXR1 del101-226 in TG1	16, 66
14	pACT2-PXR1 del150-226 in TG1	16, 67
15	pACT2-PXR1 del 226-265 in TG1	16, 68
16	pACT2-PXR1 del 7-150 in TG1	16, 69
17	pACT2-PXR1 del 150-265 in TG1	16, 70
18	pACT2-PXR1 del 101-150 (#1) in TG1	16, 71
19	pACT2-PXR1 101-226 aa #6 in TG1	17, 25
20	Ycplac111-PXR1 500bps UPS/100bps DN #1	15, 32
21	Ycplac111-PXR1delGatch #2	16, 5
22	Ycplac111-PXR1delG+PXR1G #4 in TG1	16, 6
23	Ycplac111-PXR1delG+PXR1G H55P #5 in TG1	16, 7
24	Ycplac111-PXR1delG+SPP382G #8 in TG1	16, 8
25	Ycplac111-PXR1delG+SQS1G #18 in TG1	16, 9
26	YCplac111	1, 39
27	Ycplac111-PXR1 del 101-150 (#4) in TG1 (11/9/11)	16, 72
28	Ycplac111-PXR1Δ150-226 (KKE/D) clone #5	16, 74
29	Ycplac111-PXR1 del 226-265aa #2 in TG1	17, 24
30	Ycplac111-YLR424w(6/30/03)	7, 18
31	pACT- SPP382 ΔG+SPP382G (clone #9) in TG1	14, 73
32	pACT-SPP382 ΔG+SQS1 G #5 in TG1	15, 64
33	pACT- SPP382 ΔG+PXR1G (clone #1) in TG1,	14, 80
34	Ycplac111-SPP382ΔG, clone#16 in TG1	14, 35
35	Ycplac111-SPP382ΔG+SPP382 G patch #19 in TG1	14, 44
36	Ycplac111-SPP382ΔG+PXR1 G patch #16 in	14, 47

	TG1	
37	Ycplac111-SPP382 ΔG+SQS1 G #17 in TG1	15, 63
38	pACT-Spp382delG+Pxr1 G-patch R27G clone#1	17, 50
39	pACT-Spp382delG+Pxr1 G-patch P48G clone#3	17, 51
40	pACT-SPP382ΔG+PXR1G H55P #11 in TG1	15, 31
41	pACT-Spp382delG+Pxr1 G-patch K57E clone#4	17, 52
42	pACT-Spp382delG+Pxr1 G-patch D62M clone#2	17, 53
43	Ycplac111-Spp382delG+Pxr1 G-patch R27G clone#2	17, 54
44	Ycplac111-Spp382delG+Pxr1 G-patch P48G clone#2	17, 55
45	Ycplac111-Spp382delG +PXR1 G-patch H55P #15 in TG1	15, 21
46	Ycplac111-SPP382del G + PXR1 G K57E mutant #1	15, 13
47	Ycplac111-Spp382delG+Pxr1 G-patch D62M clone#2	17, 56
48	pAS2-Prp43del NTD #3 in TG1	17, 8
49	pAS2-Prp43del RecA1 #20 in TG1	17, 9
50	pAS2-Prp43del RecA2 #4 in TG1	17, 10
51	pAS2-Prp43del WH #17 in TG1	17, 11
52	pAS2-Prp43del Ratchet #12 in TG1	17, 12
53	pAS2-Prp43del CTD #5 in TG1	17, 13
54	pAS2-Prp43 NTD #6 in DH5α	17, 15
55	pAS2-Prp43 RecA1#10 in TG1	17, 16
56	pAS2-Prp43 RecA2 #12 in TG1	17, 17
57	pAS2-Prp43 WH #1 in DH5α	17, 20
58	pAS2-Prp43 Ratchet #16 in DH5α	17, 21
59	pAS2-Prp43 CTD #1 in TG1	17, 22
60	pACT2-Prp43 in TG1 clone 1	17, 65
61	pACT2-Prp43 WH domain only in TG1 clone 11	17, 69
62	PAS2-Spp382	12, 33
63	PAS2-Sqs1	12, 31
63	pAS2-Pxr1 in TG1 clone 7	17, 67
64	pAS2-Spp382 G-patch only in TG1 clone 13	17, 59
65	pAS2-Pxr1 G-patch only in TG1 clone 1	17, 61
66	pAS2-Sqs1 G-patch only in TG1 clone 8	17, 64
67	ER2566 from NEB for protein expression	15, 47
68	ER2566 pTXB1 from NEB	15, 49
69	ER2566 pTXB1-Spp382 (amino acids 1-121,	15, 50

	WT)	
70	pTXB1-SPP382delG+PXR1G #2 in ER2566	16, 33
71	pTXB1-SPP382delG+SQS1G #5 in ER2566	16, 34
72	pTXB1-Spp382delG+PXR1G H55P #1 in ER2566 E.coli strain	16, 57

References

Aravind, L. and E. V. Koonin (1999). "G-patch: a new conserved domain in eukaryotic RNA-processing proteins and type D retroviral polyproteins." Trends Biochem Sci **24**(9): 342-344.

Arenas, J. E. and J. N. Abelson (1997). "Prp43: An RNA helicase-like factor involved in spliceosome disassembly." Proc Natl Acad Sci U S A **94**(22): 11798-11802.

Ares, M., Jr., L. Grate and M. H. Pauling (1999). "A handful of intron-containing genes produces the lion's share of yeast mRNA." RNA **5**(9): 1138-1139.

Aronova, A., D. Bacikova, L. B. Crotti, D. S. Horowitz and B. Schwer (2007). "Functional interactions between Prp8, Prp18, Slu7, and U5 snRNA during the second step of pre-mRNA splicing." RNA **13**(9): 1437-1444.

Blanton, S., A. Srinivasan and B. C. Rymond (1992). "PRP38 encodes a yeast protein required for pre-mRNA splicing and maintenance of stable U6 small nuclear RNA levels." Mol Cell Biol **12**(9): 3939-3947.

Boeke, J. D., J. Trueheart, G. Natsoulis and G. R. Fink (1987). "5-Fluoroorotic acid as a selective agent in yeast molecular genetics." Methods Enzymol **154**: 164-175.

Bohnsack, M. T., R. Martin, S. Granneman, M. Ruprecht, E. Schleiff and D. Tollervey (2009). "Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis." Mol Cell **36**(4): 583-592.

Boon, K. L., T. Auchynnikava, G. Edwalds-Gilbert, J. D. Barrass, A. P. Droop, C. Dez and J. D. Beggs (2006). "Yeast ntr1/spp382 mediates prp43 function in postspliceosomes." Mol Cell Biol **26**(16): 6016-6023.

Brow, D. A. (2002). "Allosteric cascade of spliceosome activation." Annu Rev Genet **36**: 333-360.

Burgess, S. M. and C. Guthrie (1993). "A mechanism to enhance mRNA splicing fidelity: the RNA-dependent ATPase Prp16 governs usage of a discard pathway for aberrant lariat intermediates." Cell **73**(7): 1377-1391.

Caceres, J. F., G. R. Screaton and A. R. Krainer (1998). "A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm." Genes Dev **12**(1): 55-66.

Campodonico, E. and B. Schwer (2002). "ATP-dependent remodeling of the spliceosome: intragenic suppressors of release-defective mutants of Saccharomyces cerevisiae Prp22." Genetics **160**(2): 407-415.

Chandler, S. D., A. Mayeda, J. M. Yeakley, A. R. Krainer and X. D. Fu (1997). "RNA splicing specificity determined by the coordinated action of RNA recognition motifs in SR proteins." Proc Natl Acad Sci U S A **94**(8): 3596-3601.

Chapman, K. B. and J. D. Boeke (1991). "Isolation and characterization of the gene encoding yeast debranching enzyme." Cell **65**(3): 483-492.

Chen, H. C. and S. C. Cheng (2012). "Functional roles of protein splicing factors." Biosci Rep **32**(4): 345-359.

Combs, D. J., R. J. Nagel, M. Ares, Jr. and S. W. Stevens (2006). "Prp43p is a DEAH-box spliceosome disassembly factor essential for ribosome biogenesis." Mol Cell Biol **26**(2): 523-534.

Cordin, O., J. Banroques, N. K. Tanner and P. Linder (2006). "The DEAD-box protein family of RNA helicases." Gene **367**: 17-37.

Cordin, O., D. Hahn and J. D. Beggs (2012). "Structure, function and regulation of spliceosomal RNA helicases." Curr Opin Cell Biol **24**(3): 431-438.

Crotti, L. B., D. Bacikova and D. S. Horowitz (2007). "The Prp18 protein stabilizes the interaction of both exons with the U5 snRNA during the second step of pre-mRNA splicing." Genes Dev **21**(10): 1204-1216.

Cumming, R. C., N. L. Andon, P. A. Haynes, M. Park, W. H. Fischer and D. Schubert (2004). "Protein disulfide bond formation in the cytoplasm during oxidative stress." J Biol Chem **279**(21): 21749-21758.

Dobbyn, H. C., P. A. McEwan, A. Krause, L. Novak-Frazer, J. Bella and R. T. O'Keefe (2007). "Analysis of pre-mRNA and pre-rRNA processing factor Snu13p structure and mutants." Biochem Biophys Res Commun **360**(4): 857-862.

Edgar, R. C. (2004). "MUSCLE: a multiple sequence alignment method with reduced time and space complexity." BMC Bioinformatics **5**: 113.

Edgar, R. C. (2004). "MUSCLE: multiple sequence alignment with high accuracy and high throughput." Nucleic Acids Res **32**(5): 1792-1797.

F. Sherman, G. R. F. and and J. B. Hicks (1986). "Laboratory Course Manual for Methods in Yeast Genetics." Cold Spring Harbor 1986. Cold Spring Harbor Laboratory.

Fabrizio, P., J. Dannenberg, P. Dube, B. Kastner, H. Stark, H. Urlaub and R. Luhrmann (2009). "The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome." Mol Cell **36**(4): 593-608.

Fields, S. and O. Song (1989). "A novel genetic system to detect protein-protein interactions." Nature **340**(6230): 245-246.

Frenal, K., I. Callebaut, K. Wecker, A. Prochnicka-Chalufour, N. Dendouga, S. Zinn-Justin, M. Delepierre, S. Tomavo and N. Wolff (2006). "Structural and functional characterization of the TgDRE multidomain protein, a DNA repair enzyme from Toxoplasma gondii." Biochemistry **45**(15): 4867-4874.

Gajiwala, K. S. and S. K. Burley (2000). "Winged helix proteins." Curr Opin Struct Biol **10**(1): 110-116.

Gautier, T., T. Berges, D. Tollervey and E. Hurt (1997). "Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis." Mol Cell Biol **17**(12): 7088-7098.

Gavin, A. C., P. Aloy, P. Grandi, R. Krause, M. Boesche, M. Marzioch, C. Rau, L. J. Jensen, S. Bastuck, B. Dumpelfeld, A. Edelmann, M. A. Heurtier, V. Hoffman, C. Hoefert, K. Klein, M. Hudak, A. M. Michon, M. Schelder, M. Schirle, M. Remor, T. Rudi, S. Hooper, A. Bauer, T. Bouwmeester, G. Casari, G. Drewes, G. Neubauer, J. M. Rick, B. Kuster, P. Bork, R. B. Russell and G. Superti-Furga (2006). "Proteome survey reveals modularity of the yeast cell machinery." Nature **440**(7084): 631-636.

Ghaemmaghami, S., W. K. Huh, K. Bower, R. W. Howson, A. Belle, N. Dephoure, E. K. O'Shea and J. S. Weissman (2003). "Global analysis of protein expression in yeast." Nature **425**(6959): 737-741.

Gietz, R. D. and A. Sugino (1988). "New yeast-Escherichia coli shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites." Gene **74**(2): 527-534.

Guglielmi, B. and M. Werner (2002). "The yeast homolog of human PinX1 is involved in rRNA and small nucleolar RNA maturation, not in telomere elongation inhibition." J Biol Chem **277**(38): 35712-35719.

Harper, J. W., G. R. Adami, N. Wei, K. Keyomarsi and S. J. Elledge (1993). "The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases." Cell **75**(4): 805-816.

He, Y., G. R. Andersen and K. H. Nielsen (2010). "Structural basis for the function of DEAH helicases." EMBO Rep **11**(3): 180-186.

Henras, A. K., J. Soudet, M. Gerus, S. Lebaron, M. Caizergues-Ferrer, A. Mougin and Y. Henry (2008). "The post-transcriptional steps of eukaryotic ribosome biogenesis." Cell Mol Life Sci **65**(15): 2334-2359.

Ito, H., Y. Fukuda, K. Murata and A. Kimura (1983). "Transformation of intact yeast cells treated with alkali cations." J Bacteriol **153**(1): 163-168.

James, P., J. Halladay and E. A. Craig (1996). "Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast." Genetics **144**(4): 1425-1436.

Kim, S. H. and R. J. Lin (1996). "Spliceosome activation by PRP2 ATPase prior to the first transesterification reaction of pre-mRNA splicing." Mol Cell Biol **16**(12): 6810-6819.

Kim, S. H., J. Smith, A. Claude and R. J. Lin (1992). "The purified yeast pre-mRNA splicing factor PRP2 is an RNA-dependent NTPase." EMBO J **11**(6): 2319-2326.

Kistler, A. L. and C. Guthrie (2001). "Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for sub2, an essential spliceosomal ATPase." Genes Dev **15**(1): 42-49.

Koodathingal, P., T. Novak, J. A. Piccirilli and J. P. Staley (2010). "The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA splicing." Mol Cell **39**(3): 385-395.

Kos, M. and D. Tollervey (2005). "The Putative RNA Helicase Dbp4p Is Required for Release of the U14 snoRNA from Preribosomes in Saccharomyces cerevisiae." Mol Cell **20**(1): 53-64.

Krall, J. A., E. M. Beyer and G. MacBeath (2011). "High- and low-affinity epidermal growth factor receptor-ligand interactions activate distinct signaling pathways." PLoS One **6**(1): e15945.

Kressler, D., E. Hurt and J. Bassler (2010). "Driving ribosome assembly." Biochim Biophys Acta **1803**(6): 673-683.

Leary, D. J. and S. Huang (2001). "Regulation of ribosome biogenesis within the nucleolus." FEBS Lett **509**(2): 145-150.

Lebaron, S., C. Froment, M. Fromont-Racine, J. C. Rain, B. Monsarrat, M. Caizergues-Ferrer and Y. Henry (2005). "The splicing ATPase prp43p is a component of multiple preribosomal particles." Mol Cell Biol **25**(21): 9269-9282.

Lebaron, S., C. Papin, R. Capeyrou, Y. L. Chen, C. Froment, B. Monsarrat, M. Caizergues-Ferrer, M. Grigoriev and Y. Henry (2009). "The ATPase and helicase activities of Prp43p are stimulated by the G-patch protein Pfa1p during yeast ribosome biogenesis." EMBO J **28**(24): 3808-3819.

Leeds, N. B., E. C. Small, S. L. Hiley, T. R. Hughes and J. P. Staley (2006). "The splicing factor Prp43p, a DEAH box ATPase, functions in ribosome biogenesis." Mol Cell Biol **26**(2): 513-522.

Legrain, P., B. Seraphin and M. Rosbash (1988). "Early commitment of yeast pre-mRNA to the spliceosome pathway." Mol Cell Biol **8**(9): 3755-3760.

Maeder, C. and C. Guthrie (2008). "Modifications target spliceosome dynamics." Nat Struct Mol Biol **15**(5): 426-428.

Maeder, C., A. K. Kutach and C. Guthrie (2009). "ATP-dependent unwinding of U4/U6 snRNAs by the Brr2 helicase requires the C terminus of Prp8." Nat Struct Mol Biol **16**(1): 42-48.

Martin, A., S. Schneider and B. Schwer (2002). "Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome." J Biol Chem **277**(20): 17743-17750.

Mathew, R., K. Hartmuth, S. Mohlmann, H. Urlaub, R. Ficner and R. Luhrmann (2008). "Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome." Nat Struct Mol Biol **15**(5): 435-443.

Mayas, R. M., H. Maita, D. R. Semlow and J. P. Staley (2010). "Spliceosome discards intermediates via the DEAH box ATPase Prp43p." Proc Natl Acad Sci U S A **107**(22): 10020-10025.

Mayas, R. M., H. Maita and J. P. Staley (2006). "Exon ligation is proofread by the DExD/H-box ATPase Prp22p." Nat Struct Mol Biol **13**(6): 482-490.

Miller, J. and I. Stagljar (2004). "Using the yeast two-hybrid system to identify interacting proteins." Methods Mol Biol **261**: 247-262.

Miller, J. H. (1972). "Experiments in Molecular Genetics: Assay of β-Galactosidase." CSH Laboratory Press, Cold Spring Harbor, NY: 352-355.

Mockli, N. and D. Auerbach (2004). "Quantitative beta-galactosidase assay suitable for high-throughput applications in the yeast two-hybrid system." Biotechniques **36**(5): 872-876.

Newman, A. and C. Norman (1991). "Mutations in yeast U5 snRNA alter the specificity of 5' splice-site cleavage." Cell **65**(1): 115-123.

Newman, A. J. and C. Norman (1992). "U5 snRNA interacts with exon sequences at 5' and 3' splice sites." Cell **68**(4): 743-754.

O'Day, C. L., G. Dalbadie-McFarland and J. Abelson (1996). "The Saccharomyces cerevisiae Prp5 protein has RNA-dependent ATPase activity with specificity for U2 small nuclear RNA." J Biol Chem **271**(52): 33261-33267.

Ostergaard, H., C. Tachibana and J. R. Winther (2004). "Monitoring disulfide bond formation in the eukaryotic cytosol." J Cell Biol **166**(3): 337-345.

Palle, K., L. Pattarello, M. van der Merwe, C. Losasso, P. Benedetti and M. A. Bjornsti (2008). "Disulfide cross-links reveal conserved features of DNA topoisomerase I architecture and a role for the N terminus in clamp closure." J Biol Chem **283**(41): 27767-27775.

Pandit, S., B. Lynn and B. C. Rymond (2006). "Inhibition of a spliceosome turnover pathway suppresses splicing defects." Proc Natl Acad Sci U S A **103**(37): 13700-13705.

Pandit, S., S. Paul, L. Zhang, M. Chen, N. Durbin, S. M. Harrison and B. C. Rymond (2009). "Spp382p interacts with multiple yeast splicing factors, including possible regulators of Prp43 DExD/H-Box protein function." Genetics **183**(1): 195-206.

Pandit, S. S. P. L., Zhang; M.Chen; N.Durbin; S. Harrison and B.C.Rymond (2009). "Spp382p promotes intracellular spliceosome recycling and interacts with a diverse set of factors including an inhibitor of Prp43 DExD/H protein function. ." Submitted.

Perriman, R., I. Barta, G. K. Voeltz, J. Abelson and M. Ares, Jr. (2003). "ATP requirement for Prp5p function is determined by Cus2p and the structure of U2 small nuclear RNA." Proc Natl Acad Sci U S A **100**(24): 13857-13862.

Pertschy, B., C. Schneider, M. Gnadig, T. Schafer, D. Tollervey and E. Hurt (2009). "RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1." J Biol Chem **284**(50): 35079-35091.

Pleiss, J. A., G. B. Whitworth, M. Bergkessel and C. Guthrie (2007). "Rapid, transcript-specific changes in splicing in response to environmental stress." Mol Cell **27**(6): 928-937.

Plocik, A. M. and C. Guthrie (2012). "Diverse forms of RPS9 splicing are part of an evolving autoregulatory circuit." PLoS Genet **8**(3): e1002620.

Raghunathan, P. L. and C. Guthrie (1998). "RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2." Curr Biol **8**(15): 847-855.

Rocak, S. and P. Linder (2004). "DEAD-box proteins: the driving forces behind RNA metabolism." Nat Rev Mol Cell Biol **5**(3): 232-241.

Rymond, B. C., C. Pikielny, B. Seraphin, P. Legrain and M. Rosbash (1990). "Measurement and analysis of yeast pre-mRNA sequence contribution to splicing efficiency." Methods Enzymol **181**: 122-147.

Sambrook J, Fritsch EF and M. T. (1989). "Molecular Cloning, A Laboratory Manual, 2nd Edition."

Schneider, S., E. Campodonico and B. Schwer (2004). "Motifs IV and V in the DEAH box splicing factor Prp22 are important for RNA unwinding, and helicase-defective Prp22 mutants are suppressed by Prp8." J Biol Chem **279**(10): 8617-8626.

Schneider, S., H. R. Hotz and B. Schwer (2002). "Characterization of dominant-negative mutants of the DEAH-box splicing factors Prp22 and Prp16." J Biol Chem **277**(18): 15452-15458.

Schwer, B. (2001). "A new twist on RNA helicases: DExH/D box proteins as RNPases." Nat Struct Biol **8**(2): 113-116.

Schwer, B. (2008). "A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release." Mol Cell **30**(6): 743-754.

Schwer, B. and C. H. Gross (1998). "Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing." EMBO J **17**(7): 2086-2094.

Schwer, B. and C. Guthrie (1991). "PRP16 is an RNA-dependent ATPase that interacts transiently with the spliceosome." Nature **349**(6309): 494-499.

Sikorski, R. S. and J. D. Boeke (1991). "In vitro mutagenesis and plasmid shuffling: from cloned gene to mutant yeast." Methods Enzymol **194**: 302-318.

Silverman, E. J., A. Maeda, J. Wei, P. Smith, J. D. Beggs and R. J. Lin (2004). "Interaction between a G-patch protein and a spliceosomal DEXD/H-box ATPase that is critical for splicing." Mol Cell Biol **24**(23): 10101-10110.

Sonnhammer, E. L. and V. Hollich (2005). "Scoredist: a simple and robust protein sequence distance estimator." BMC Bioinformatics **6**: 108.

Spingola, M., L. Grate, D. Haussler and M. Ares, Jr. (1999). "Genome-wide bioinformatic and molecular analysis of introns in Saccharomyces cerevisiae." RNA **5**(2): 221-234.

Staley, J. P. and C. Guthrie (1999). "An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p." Mol Cell **3**(1): 55-64.

Stevens, S. W., D. E. Ryan, H. Y. Ge, R. E. Moore, M. K. Young, T. D. Lee and J. Abelson (2002). "Composition and functional characterization of the yeast spliceosomal penta-snRNP." Mol Cell **9**(1): 31-44.

Strauss, E. J. and C. Guthrie (1994). "PRP28, a 'DEAD-box' protein, is required for the first step of mRNA splicing in vitro." Nucleic Acids Res **22**(15): 3187-3193.

Tanaka, N., A. Aronova and B. Schwer (2007). "Ntr1 activates the Prp43 helicase to trigger release of lariat-intron from the spliceosome." Genes Dev **21**(18): 2312-2325.

Tanaka, N. and B. Schwer (2005). "Characterization of the NTPase, RNA-binding, and RNA helicase activities of the DEAH-box splicing factor Prp22." Biochemistry **44**(28): 9795-9803.

Tardiff, D. F. and M. Rosbash (2006). "Arrested yeast splicing complexes indicate stepwise snRNP recruitment during in vivo spliceosome assembly." RNA **12**(6): 968-979.

Tsai, R. T., R. H. Fu, F. L. Yeh, C. K. Tseng, Y. C. Lin, Y. H. Huang and S. C. Cheng (2005). "Spliceosome disassembly catalyzed by Prp43 and its associated components Ntr1 and Ntr2." Genes Dev **19**(24): 2991-3003.

Umen, J. G. and C. Guthrie (1995). "The second catalytic step of pre-mRNA splicing." RNA **1**(9): 869-885.

Valadkhan, S. (2005). "snRNAs as the catalysts of pre-mRNA splicing." Curr Opin Chem Biol **9**(6): 603-608.

van Der Houven Van Oordt, W., K. Newton, G. R. Screaton and J. F. Caceres (2000). "Role of SR protein modular domains in alternative splicing specificity in vivo." Nucleic Acids Res **28**(24): 4822-4831.

Wach, A., A. Brachat, R. Pohlmann and P. Philippsen (1994). "New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae." Yeast **10**(13): 1793-1808.

Wagner, J. D., E. Jankowsky, M. Company, A. M. Pyle and J. N. Abelson (1998). "The DEAH-box protein PRP22 is an ATPase that mediates ATP-dependent mRNA release from the spliceosome and unwinds RNA duplexes." EMBO J **17**(10): 2926-2937.

Walbott, H., S. Mouffok, R. Capeyrou, S. Lebaron, O. Humbert, H. van Tilbeurgh, Y. Henry and N. Leulliot (2010). "Prp43p contains a processive helicase structural architecture with a specific regulatory domain." EMBO J **29**(13): 2194-2204.

Wang, Q., L. Zhang, B. Lynn and B. C. Rymond (2008). "A BBP-Mud2p heterodimer mediates branchpoint recognition and influences splicing substrate abundance in budding yeast." Nucleic Acids Res **36**(8): 2787-2798.

Wang, Y., J. D. Wagner and C. Guthrie (1998). "The DEAH-box splicing factor Prp16 unwinds RNA duplexes in vitro." Curr Biol **8**(8): 441-451.

Warner, J. R. (1999). "The economics of ribosome biosynthesis in yeast." Trends Biochem Sci **24**(11): 437-440.

Will, C. L. and R. Luhrmann (2011). "Spliceosome structure and function." Cold Spring Harb Perspect Biol **3**(7).

Winzeler, E. A., D. D. Shoemaker, A. Astromoff, H. Liang, K. Anderson, B. Andre, R. Bangham, R. Benito, J. D. Boeke, H. Bussey, A. M. Chu, C. Connelly, K. Davis, F. Dietrich, S. W. Dow, M. El Bakkoury, F. Foury, S. H. Friend, E. Gentalen, G. Giaever, J. H. Hegemann, T. Jones, M. Laub, H. Liao, N. Liebundguth, D. J. Lockhart, A. Lucau-Danila, M. Lussier, N. M'Rabet, P. Menard, M. Mittmann, C. Pai, C. Rebischung, J. L. Revuelta, L. Riles, C. J. Roberts, P. Ross-MacDonald, B. Scherens, M. Snyder, S. Sookhai-Mahadeo, R. K. Storms, S. Veronneau, M. Voet, G. Volckaert, T. R. Ward, R. Wysocki, G. S. Yen, K. Yu, K. Zimmermann, P. Philippsen, M. Johnston and R. W.

Davis (1999). "Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis." Science **285**(5429): 901-906.

Woodman, I. L. and E. L. Bolt (2011). "Winged helix domains with unknown function in Hel308 and related helicases." Biochem Soc Trans **39**(1): 140-144.

Xie, J., K. Beickman, E. Otte and B. C. Rymond (1998). "Progression through the spliceosome cycle requires Prp38p function for U4/U6 snRNA dissociation." EMBO J **17**(10): 2938-2946.

VITA

Daipayan Banerjee

Born: 1982, Berhampore, Murshidabad District, West Bengal, INDIA.

Education

- University of Kentucky (2006-Present)

  Doctoral Program (Biology)

  Specilaization: Molecular Biology

- West Bengal University of Technology.

  Bengal College of Engineering & Technology, Durgapur, India

  Bachelor in Technology in Biotechnology (2001-2005)

Awards

- Outstanding Teaching Assistant Award, College of Arts & Sciences, University of Kentucky, 2011
- Ribble Mini Grant, Department of Biology, University of Kentucky, 2011
- Ribble Mini Grant, Department of Biology, University of Kentucky, 2010